



QY	301	AACTTTGCAATGGAAAAATTTGAATGATGGGATTTATCAACGGAAATTCGAAAAATATAA	360
Db	301	AACTTTGCAATGGAAAAATTTGAATGATGGGATTTATCAACGGAAATTCGAAAAATATAA	360
QY	361	ATTGGAGGAAATTTATTTCTTAAGCTCGGAAGGGAATGGCAGTAGAAAAATGAGATT	420
Db	361	ATTGGAGGAAATTTATTTCTTAAGCTCGGAAGGGAATGGCAGTAGAAAAATGAGATT	420
QY	421	ATCAATGCTGGTTCTTTTCATTCTATATATCCAAAAAAGATGATTTTAAAGGCTTTG	480
Db	421	ATCAATGCTGGTTCTTTTCATTCTATATATCCAAAAAAGATGATTTTAAAGGCTTTG	480
QY	481	GAGAGAGCCAAACATGATAAGTTTAAATGGAATCATTCGAGTAGATGAAAATGATAAA	540
Db	481	GAGAGAGCCAAACATGATAAGTTTAAATGGAATCATTCGAGTAGATGAAAATGATAAA	540
QY	541	ATTCATTGAATCCGATGGAAAGCATTCGATGAGAAAGGAAAAATCATGCTTTGAAAGC	600
Db	541	ATTCATTGAATCCGATGGAAAGCATTCGATGAGAAAGGAAAAATCATGCTTTGAAAGC	600
QY	601	ATCGGTTATATGCGCGCGATTTTGATTGAAAGATCTGCATATCTAAAGCAGGAATT	660
Db	601	ATCGGTTATATGCGCGCGATTTTGATTGAAAGATCTGCATATCTAAAGCAGGAATT	660
QY	661	ACAGATTTTAAAAATTTAGTCAATTTAGTGAATTCGAATTAATTCGATCGCGAGAT	720
Db	661	ACGATTTTAAAAATTTAGTCAATTTAGTGAATTCGAATTAATTCGATCGCGAGAT	720
QY	721	TTAAAGCTACCAAGACAAATCTGGAATATTAATCTTTCAGCTCATAGATTCTCT	780
Db	721	TTAAAGCTACCAAGACAAATCTGGAATATTAATCTTTCAGCTCATAGATTCTCTCT	780
QY	781	CAAAAGCTATGGGAAAAATTCACCTGTGAAAAGAAATGAAATATGTAAAGGA	840
Db	781	CAAAAGCTATGGGAAAAATTCACCTGTGAAAAGAAATGAAATATGTAAAGGA	840
QY	841	AATACCAAGCAAAATATTGAATCTGATGCTGTATTGAAAGCAGATGAAAATATAAAATT	900
Db	841	AATACCAAGCAAAATATTGAATCTGATGCTGTATTGAAAGCAGATGAAAATATAAAATT	900
QY	901	AGTGGAAAGCTACAAATGGGAGATTTATTAAGAAAGAAAGGGGAAAAAGAACTTATAC	960
Db	901	AGTGGAAAGCTACAAATGGGAGATTTATTAAGAAAGAAAGGGGAAAAAGAACTTATAC	960
QY	961	ACTCCTTAAAGTTATCGAATGGAAGCTTCGTAAGATTAATAAGGAAAACCTCAT	1020
Db	961	ACTCCTTAAAGTTATCGAATGGAAGCTTCGTAAGATTAATAAGGAAAACCTCAT	1020
QY	1021	GGAAGAAATGTTGACATTCACAGCTGAAGCAAGAAATTTCTATGATGCAACTTACT	1080
Db	1021	GGAAGAAATGTTGACATTCACAGCTGAAGCAAGAAATTTCTATGATGCAACTTACT	1080
QY	1081	AAAGCTTGCAAGCACTCTTTTACGTTTGTACAGGTTCTATTTCTCTATCAATTTAAAT	1140
Db	1081	AAAGCTTGCAAGCACTCTTTTACGTTTGTACAGGTTCTATTTCTCTATCAATTTAAAT	1140
QY	1141	GGAATTTTAAAGTTTATGCAATGATGACAGTGCATGTTATGGAAGAAATGGCCAAAGTC	1200
Db	1141	GGAATTTTAAAGTTTATGCAATGATGACAGTGCATGTTATGGAAGAAATGGCCAAAGTC	1200
QY	1201	GAAAGCAACAGAAAGGCAATATTCATTTCTTACAGTGAAGTAAAGCACTATGGGA	1260
Db	1201	GAAAGCAACAGAAAGGCAATATTCATTTCTTACAGTGAAGTAAAGCACTATGGGA	1260
QY	1261	GAGAGCTACTTCTCCATTAATAATTAACAAATTTATTTGGAGAAAGCCCAATGGAAACTT	1320
Db	1261	GAGAGCTACTTCTCCATTAATAATTAACAAATTTATTTGGAGAAAGCCCAATGGAAACTT	1320
QY	1321	CTCAGTATCGAGCGGAGATATATTTCTGCAAAAAGTAATTCAAATGTATCTATTAAGGA	1380
Db	1321	CTCAGTATCGAGCGGAGATATATTTCTGCAAAAAGTAATTCAAATGTATCTATTAAGGA	1380

QY	1381	GAAGTAAATCGAAGGGAAGGACGATATTTCTTCAAAAATCGAAAAATACATTAAGTCT	14410
Db	1391	GAAGTAAATCGAAGGGAAGGACGATATTTCTTCAAAAATCGAAAAATACATTAAGTCT	14400
QY	1441	TCTGTTCTGTGGAAACGATGAGATTCATATAAGAGCTCTTTCAGTATTGGTAGC	15000
Db	1441	TCTGTTCTGTGGAAACGATGAGATTCATATAAGAGCTCTTTCAGTATTGGTAGC	15000
QY	1501	GAAGGAAAAATTAATCTTCCGTCAGAATTGCTAAAGGACGAAAAGTGAATCGAAAAAG	15600
Db	1501	GAAGGAAAAATTAATCTTCCGTCAGAATTGCTAAAGGACGAAAAGTGAATCGAAAAAG	15600
QY	1551	GATATATGTAATTCGAGAAAGTGAAGCCATTAAATTCATTCGAGCTCGTATAAAGTGG	16200
Db	1551	GATATATGTAATTCGAGAAAGTGAAGCCATTAAATTCATTCGAGCTCGTATAAAGTGG	16200
QY	1621	TTGGGGGATAGTGGTATATGGGGTTGTGGCTGCAATATTTCTAATTAATAGTCTCTCC	16880
Db	1621	TTGGGGGATAGTGGTATATGGGGTTGTGGCTGCAATATTTCTAATTAATAGTCTCTCC	16880
QY	1661	CGTATAGATGTAGATGAGTATCTACATGCGAAGGCGCTAAATGTGGAGCTCATTAAC	17400
Db	1661	CGTATAGATGTAGATGAGTATCTACATGCGAAGGCGCTAAATGTGGAGCTCATTAAC	17400
QY	1741	ATTACTAAAAATAGTGTCTGCAACAGAGCTGATTTTGGGAACTTCCAACTTTATGAT	18000
Db	1741	ATTACTAAAAATAGTGTCTGCAACAGAGCTGATTTTGGGAACTTCCAACTTTATGAT	18000
QY	1801	GATCAGCTTTATGATCAGGTCACTCTAAATCAATTTTATGATGCAATMAACAGCGGTTT	18600
Db	1801	GATCAGCTTTATGATCAGGTCACTCTAAATCAATTTTATGATGCAATMAACAGCGGTTT	18600
QY	1861	GGAAGAGACAGTGTCAATGAGAAAAATAAGATAAGCTTAACGAACTTATTTAGTCCGT	19200
Db	1861	GGAAGAGACAGTGTCAATGAGAAAAATAAGATAAGCTTAACGAACTTATTTAGTCCGT	19200
QY	1921	GTCGTCGAACCATAGCAAAATCATTAATTTCTGCTCTGTGGCATTAAGAGAGATGGA	19880
Db	1921	GTCGTCGAACCATAGCAAAATCATTAATTTCTGCTCTGTGGCATTAAGAGAGATGGA	19880
QY	1981	AGACTTTCTTCAGAGTGGAGAGGAGTATATTAAGGGCATTAATTAAGAGCTCAAAATCTT	20400
Db	1981	AGACTTTCTTCAGAGTGGAGAGGAGTATATTAAGGGCATTAATTAAGAGCTCAAAATCTT	20400
QY	2041	CGAGGCACTACGTCAAGTGAAGTGTGGCTGTACGAAAGGAGAAAAAAGAAAACTTAT	21000
Db	2041	CGAGGCACTACGTCAAGTGAAGTGTGGCTGTACGAAAGGAGAAAAAAGAAAACTTAT	21000
QY	2101	GGAATATGACAGATTTTATATGAAACTATTAATAATTAATGCTTCTGTGCAATTTGCCAT	21600
Db	2101	GGAATATGACAGATTTTATATGAAACTATTAATAATTAATGCTTCTGTGCAATTTGCCAT	21600
QY	2161	CATGCTGAATTTGGTATGCGAAGGAAAAATTAATTAATCAACAGTGAATAATTAATTAAT	22200
Db	2161	CATGCTGAATTTGGTATGCGAAGGAAAAATTAATTAATCAACAGTGAATAATTAATTAAT	22200
QY	2221	AAAAATCCTTCAAAAATTCGCAAGTCTGTATTTGATTAATTAAGACCTTTTAAGAGAGCT	22880
Db	2221	AAAAATCCTTCAAAAATTCGCAAGTCTGTATTTGATTAATTAAGACCTTTTAAGAGAGCT	22880
QY	2281	TTTGGAAAAAGAACGAAAACTCCAGAAATATGATCCGAAAGATTTTGAATCTATTGAAAA	23400
Db	2281	TTTGGAAAAAGAACGAAAACTCCAGAAATATGATCCGAAAGATTTTGAATCTATTGAAAA	23400
QY	2341	TTATTTGAATGATTTTTCGAAAAATTTGATGAAAAACCGAGCTTTTATCTAAATGGTGA	24000
Db	2341	TTATTTGAATGATTTTTCGAAAAATTTGATGAAAAACCGAGCTTTTATCTAAATGGTGA	24000
QY	2401	AGAAATGACAAATTAATTTCTCCGAGTGAACCTCAAAAACAGAACTGCTATTGAATTTGCA	24600
Db	2401	AGAAATGACAAATTAATTTCTCCGAGTGAACCTCAAAAACAGAACTGCTATTGAATTTGCA	24600
QY	2461	AACATATGTCAGGAGAAATGAAAAATTTAGAGAAAAATTAACGAAAGATTTAAAGCT	25200

Dp	2461	AACTATGTTGACGGAGAAATGAAAAAATTTGAGGAAAAATTTACCGAAAGGATTTAAAGCT	2520
Qy	2521	TTTTCAAGAAATGATGATGACCTGATTAAGAAAATTGAAATTTTACAGAGTAGGAAAT	2580
Dp	2521	TTTTCAAGAGATTTGATGATGACTGATTTAAAGAACTTTGAATTTTACAGAGTAGGAAAT	2580
Qy	2581	TATGCAAAATTTTCCACACTTTTCTCTCTCCGAGCTAATGAGAAAAGATGTTTCTTCT	2640
Dp	2581	TATGCAAAATTTTCCACACTTTTCTCTCTCCGAGCTAATGAGAAAAGATGTTTCTTCT	2640
Qy	2641	GTGGAGAGAGCTGTTTGTGGGTGAAACAGAGAAATTATGCAAGTATCCGTTGGAAAA	2700
Dp	2641	GTGGAGAGAGCTGTTTGTGGGTGAAACAGAGAAATTATGCAAGTATCCGTTGGAAAA	2700
Qy	2701	GGAGCTAAACTGCTGCGAAAAAAAGATTTTAAATTAAGCTATCATTAAGCGAAGACA	2760
Dp	2701	GGAGCTAAACTGCTGCGAAAAAAAGATTTTAAATTAAGCTATCATTAAGCGAAGACA	2760
Qy	2761	GTGAATTTAGTTGGAAAAATTGGACTTGCAGAAAGCAATATCCGGAAGTGCAGTCGA	2820
Dp	2761	GTGAATTTAGTTGGAAAAATTGGACTTGCAGAAAGCAATATCCGGAAGTGCAGTCGA	2820
Qy	2821	GGAAATTTAATGTTTCAAAAGATCGAAAAATTCAGCTATCGTAGAAGCTTAAAGAAAAGCT	2880
Dp	2821	GGAAATTTAATGTTTCAAAAGATCGAAAAATTCAGCTATCGTAGAAGCTTAAAGAAAAGCT	2880
Qy	2881	GAATTTATCAGAGAAAAATTAATGACATGCAATTTGAAACAGACTTTTCAATGAGCGGA	2940
Dp	2881	GAATTTATCAGAGAAAAATTAATGACATGCAATTTGAAACAGACTTTTCAATGAGCGGA	2940
Qy	2941	TCTTTTAATGCTGCTCAGGTGGGAATGCAATCAATGGAATGGGAATTAATGAGAGGT	3000
Dp	2941	TCTTTTAATGCTGCTCAGGTGGGAATGCAATCAATGGAATGGGAATTAATGAGAGGT	3000
Qy	3001	ATCAGTAAAGCAAGATTTCCATTGATGACGAGCAATATTTGAAAGCTAATTAAAAAATT	3060
Dp	3001	ATCAGTAAAGCAAGATTTCCATTGATGACGAGCAATATTTGAAAGCTAATTAAAAAATT	3060
Qy	3061	GCTTTAAACATTAAGAAATGATATCTTCGTTTGGAAATGCTGCGGTTACAGCGGGAATCGGA	3120
Dp	3061	GCTTTAAACATTAAGAAATGATATCTTCGTTTGGAAATGCTGCGGTTACAGCGGGAATCGGA	3120
Qy	3121	ACGAAAAATGCGCGCTGCGGGTTCGTTGCTGCTGCGTAAATGATTAATTTCAACAAA	3180
Dp	3121	ACGAAAAATGCGCGCTGCGGGTTCGTTGCTGCTGCGTAAATGATTAATTTCAACAAA	3180
Qy	3181	GCTTCATTGAAAGATATGACGAGAGCAAAAGTAAATATGATTAAGAAATTAAGATGATGA	3240
Dp	3181	GCTTCATTGAAAGATATGACGAGAGCAAAAGTAAATATGATTAAGAAATTAAGATGATGA	3240
Qy	3241	GTAACAGTAACTGCGCGAATCTTTTGAAGTAAATGCAAAAACGACCGGAACATTAACAGT	3300
Dp	3241	GTAACAGTAACTGCGCGAATCTTTTGAAGTAAATGCAAAAACGACCGGAACATTAACAGT	3300
Qy	3301	AATTCGTGTCGCGAGGAATTAATPAAGTTGGAAGTAAACGAGTGAAGAAAAACGAAA	3360
Dp	3301	AATTCGTGTCGCGAGGAATTAATPAAGTTGGAAGTAAACGAGTGAAGAAAAACGAAA	3360
Qy	3361	TCAGAAAGAAAGCCAGAGGGAATTTTGTGCAAAATCGGAAACAAAGTGAATCTGTGTA	3420
Dp	3361	TCAGAAAGAAAGCCAGAGGGAATTTTGTGCAAAATCGGAAACAAAGTGAATCTGTGTA	3420
Qy	3421	AATAAATTTACGAGATGATGATTCATTAACAGAAAAAATTACAAATTAATTTCTGAA	3480
Dp	3421	AATAAATTTACGAGATGATGATTCATTAACAGAAAAAATTACAAATTAATTTCTGAA	3480
Qy	3481	GGAGTAAAGAAAAAGCGGGGAATTTCTCTCGAAAGTTTCTCATCTCCCGATTAAGAGACG	3540
Dp	3481	GGAGTAAAGAAAAAGCGGGGAATTTCTCTCGAAAGTTTCTCATCTCCCGATTAAGAGACG	3540
Qy	3541	TCTTTCAAGTTGGAGCTTCTGGAAGTGTCTTTCAATTAATTTAAAGAAACATCT	3600

Db	3541	TCCTTCAGCTTTGGGAGCTTCTCGAAAGTGTTCTTTCAATAAATTTAAAAAGAAACAATCT	3600
Qy	3601	GCTGTCGTAGATGGAGTAAAGATAAATTTGAAGGAGCAAAATATAAAAGGTAGAGTGACT	3660
Db	3601	GCTGTCGTAGATGGAGTAAAGATAAATTTGAAGGAGCAAAATATAAAAGGTAGAGTGACT	3660
Qy	3661	TCCTTCGATTCCTACTTTTGTTGGAGCATGGGGCGGAGCTGCTGCACCTTCAGTGGAAATCAT	3720
Db	3661	TCCTTCGATTCCTACTTTTGTTGGAGCATGGGGCGGAGCTGCTGCACCTTCAGTGGAAATCAT	3720
Qy	3721	ATTGGAAGTGAATATAGCAACATCAGAGCTGTTTACCTGAGCGGCTGCTGTAAATAT	3780
Db	3721	ATTGGAAGTGAATATAGCAACATCAGAGCTGTTTACTGGAAGCGGCTGCTGTAAATAT	3780
Qy	3781	ATTCAAGATAAACAAGTGCTTTGGTTAAAAATATGTATATTCGAAATGCCAATAATTT	3840
Db	3781	ATTCAAGATAAACAAGTGCTTTGGTTAAAAATATGTATATTCGAAATGCCAATAATTT	3840
Qy	3841	AAAGTAAATGCTTTGAGATGGAGAACTCAAGTACACAGAGAGAGGTTTGGAAAGAGTT	3900
Db	3841	AAAGTAAATGCTTTGAGATGGAGAACTCAAGTACACAGAGAGAGGTTTGGAAAGAGTT	3900
Qy	3901	AAAGAAGTGAAGACAAAGAAAAAGTTATCTATTGGGAACTTTCGCTCTATCAACTTA	3960
Db	3901	AAAGAAGTGAAGACAAAGAAAAAGTTATCTATTGGGAACTTTCGCTCTATCAACTTA	3960
Qy	3961	GTGAACATGAAAGTCTTCGCAAAATCAGAAAAATATAACAGTACAGAGAAATCTGAAGC	4020
Db	3961	GTGAACATGAAAGTCTTCGCAAAATCAGAAAAATATAACAGTACAGAGAAATCTGAAGC	4020
Qy	4021	CAAAAAATGATGTTGATGTCACTGCTTATCAAGCGGACACCCAAAGTACAGAGAGCTTTA	4080
Db	4021	CAAAAAATGATGTTGATGTCACTGCTTATCAAGCGGACACCCAAAGTACAGAGAGCTTTA	4080
Qy	4081	AATTTACAAGCTGGAAGTCAATGGAACGTGTAGGGGCTACTGACCTGTGGCCAAATTA	4140
Db	4081	AATTTACAAGCTGGAAGTCAATGGAACGTGTAGGGGCTACTGACCTGTGGCCAAATTA	4140
Qy	4141	AACAACAAGTAAATGCTTCTATTAGTGTGGAGATATATCTAAAGTTATCGACGGAC	4200
Db	4141	AACAACAAGTAAATGCTTCTATTAGTGTGGAGATATATCTAAAGTTATCGACGGAC	4200
Qy	4201	GCAAAAGCTCTTTTAGCAACCACTCAAGTGCCTCTGACGAGACGCGGAGGACAAAT	4260
Db	4201	GCAAAAGCTCTTTTAGCAACCACTCAAGTGCCTCTGACGAGACGCGGAGGACAAAT	4260
Qy	4261	AGTCTGAGGGGGATTAGGAAATTAATCAAGGGGCTGTTTGTCAATTAAGTTGACAAAT	4320
Db	4261	AGTCTGAGGGGGATTAGGAAATTAATCAAGGGGCTGTTTGTCAATTAAGTTGACAAAT	4320
Qy	4321	GACGTGGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAAATGAATCAATGCTATT	4380
Db	4321	GACGTGGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAAATGAATCAATGCTATT	4380
Qy	4381	GCCAAAGATGTCAAAAGAAAGTTCTGATCTAGCAAAAAGAAATTCAGGCTTTACTAAATGGA	4440
Db	4381	GCCAAAGATGTCAAAAGAAAGTTCTGATCTAGCAAAAAGAAATTCAGGCTTTACTAAATGGA	4440
Qy	4441	AAAGTAAATAATTTTGAAGATCGTGTATTAATACGACTGGAAATGGTTATTAATACG	4500
Db	4441	AAAGTAAATAATTTTGAAGATCGTGTATTAATACGACTGGAAATGGTTATTAATACG	4500
Qy	4501	AAGGAACAATAGAAAAGCAAAAGAAAAAGAGAGCGGTCAATTGTAAATGCTGCTTTA	4560
Db	4501	AAGGAACAATAGAAAAGCAAAAGAAAAAGAGAGCGGTCAATTGTAAATGCTGCTTTA	4560
Qy	4561	TCGGTTGCTGGAACGGATTAATCCGCTGAGAGAGTACTAATTGCACTCAATACGTGTTAA	4620
Db	4561	TCGGTTGCTGGAACGGATTAATCCGCTGAGAGAGTACTAATTGCACTCAATACGTGTTAA	4620
Qy	4621	AATAAATTTAAAGCAAAATGAGTGGAGCAATTAAGAAAGCCGAGAGAGATATAATTCAT	4680
Db	4621	AATAAATTTAAAGCAAAATGAGTGGAGCAATTAAGAAAGCCGAGAGAGATATAATTCAT	4680

QY 4681 GCGAAACATGTAAATGTGAGAGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGGA 4740  
DB 4681 GCGAAACATGTAAATGTGAGAGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGGA 4740  
QY 4741 CTTCCTATCAGCAAAAGATGCTTTTTCAGGAATGGATCTGAGCATGCGAAGACTTATCA 4800  
DB 4741 CTTCCTATCAGCAAAAGATGCTTTTTCAGGAATGGATCTGAGCATGCGAAGACTTATCA 4800  
QY 4801 AATGACAGATTGCAAGGTGGAATAAGGAGAAATTTCTGCTGATTCCTTAAATGCTGAC 4860  
DB 4801 AATGACAGATTGCAAGGTGGAATAAGGAGAAATTTCTGCTGATTCCTTAAATGCTGAC 4860  
QY 4861 GCAAAATTAATTCATTCTTGCGGTGATGTTGCGGAAACCATGCGGTTCTCTTCTACG 4920  
DB 4861 GCAAAATTAATTCATTCTTGCGGTGATGTTGCGGAAACCATGCGGTTCTCTTCTACG 4920  
QY 4921 GCGGTAGAGACTGCTTTTTCGGAATAATCTCTTCAATATAAAACCTGCTTGTATACA 4980  
DB 4921 GCGGTAGAGACTGCTTTTTCGGAATAATCTCTTCAATATAAAACCTGCTTGTATACA 4980  
QY 4981 GGAACGAAGTAAATCCTTTTAGTGAAGAAATAGAAAGTCAATGTAACAAGCTTGAAT 5040  
DB 4981 GGAACGAAGTAAATCCTTTTAGTGAAGAAATAGAAAGTCAATGTAACAAGCTTGAAT 5040  
QY 5041 GATTCTCATATTACAAACGTTTCTGCTGAGAGCGCTGCAAGTATTAGCAGGCTGGAATC 5100  
DB 5041 GATTCTCATATTACAAACGTTTCTGCTGAGAGCGCTGCAAGTATTAGCAGGCTGGAATC 5100  
QY 5101 GGAGGAATGATCTGTGCAATGTGTGTTCTGATGAAAGGAAAGCTTTAGTTAGTATCT 5160  
DB 5101 GGAGGAATGATCTGTGCAATGTGTGTTCTGATGAAAGGAAAGCTTTAGTTAGTATCT 5160  
QY 5161 GAGTTTGAAGAGTAAAGTCTTCTCAATGATGAGCAAAAGATCAAAAACAAATAATACA 5220  
DB 5161 GAGTTTGAAGAGTAAAGTCTTCTCAATGATGAGCAAAAGATCAAAAACAAATAATACA 5220  
QY 5221 ATTGCGGAAATGCAAAATGAGAGAAAGCGCTGAGTTGAGCAACAGTTGCTCATACA 5280  
DB 5221 ATTGCGGAAATGCAAAATGAGAGAAAGCGCTGAGTTGAGCAACAGTTGCTCATACA 5280  
QY 5281 AATATTGGAAAACAATCAGTTATAGCTATTTGTAAAAAACAGTAATAATTACAAACGCGGAT 5340  
DB 5281 AATATTGGAAAACAATCAGTTATAGCTATTTGTAAAAAACAGTAATAATTACAAACGCGGAT 5340  
QY 5341 GATCAAGATAGAAAAATATCAATGTAAGTGAAGTAATATACATATGCAATCTATTA 5400  
DB 5341 GATCAAGATAGAAAAATATCAATGTAAGTGAAGTAATATACATATGCAATCTATTA 5400  
QY 5401 GCAGTCGAGTTGAGAGCAAAAGAGCTCTGTCGAAAGAGCTTTCGAAAGTACTAC 5460  
DB 5401 GCAGTCGAGTTGAGAGCAAAAGAGCTCTGTCGAAAGAGCTTTCGAAAGTACTAC 5460  
QY 5461 TTGGAATTAAGACAGTTTCTTCATGTTGATCAAACTGATATTGACAAAAGATTTAAGGAA 5520  
DB 5461 TTGGAATTAAGACAGTTTCTTCATGTTGATCAAACTGATATTGACAAAAGATTTAAGGAA 5520  
QY 5521 GAAATTAATGAAATTAAGGAAAGGCAAAATGTTAATGTTCTAGCTGAAATTCGAGTGA 5580  
DB 5521 GAAATTAATGAAATTAAGGAAAGGCAAAATGTTAATGTTCTAGCTGAAATTCGAGTGA 5580  
QY 5581 GTGTCACAAATGCAAGTGTCTTTCGAGAGCAATGAGCAAGCTGCAAGTGAAGTCTGA 5640  
DB 5581 GTGTCACAAATGCAAGTGTCTTTCGAGAGCAATGAGCAAGCTGCAAGTGAAGTCTGA 5640  
QY 5641 GTAGCAGTTAATTAATTAACAAAAATCTTTCGACATATAAATAATAGTCTCAAAAT 5700  
DB 5641 GTAGCAGTTAATTAATTAACAAAAATCTTTCGACATATAAATAATAGTCTCAAAAT 5700  
QY 5701 GTAGCAAAATGCTTGTGTAAAGCAAAATCTCATCTATTAATAAATGGAATTGGA 5760  
DB 5701 GTAGCAAAATGCTTGTGTAAAGCAAAATCTCATCTATTAATAAATGGAATTGGA 5760

QY 5761 GCTGAGATTGAGCTGAGAGAGCTGAGATGACAGGTTCTGTAGCAGTGAATTAAGATTGTA 5820  
DB 5761 GCTGAGATTGAGCTGAGAGAGCTGAGATGACAGGTTCTGTAGCAGTGAATTAAGATTGTA 5820  
QY 5821 AATAATACGATAGCAAAATTAATATGCAAAAAATCACTGCGAAAGGAAATGTGGAATT 5880  
DB 5821 AATAATACGATAGCAAAATTAATATGCAAAAAATCACTGCGAAAGGAAATGTGGAATT 5880  
QY 5881 ATTACAGAGCTGATGCGGTAAATGCTTAATATAGCAGAAACAGTGTGAGTGCGCCGT 5940  
DB 5881 ATTACAGAGCTGATGCGGTAAATGCTTAATATAGCAGAAACAGTGTGAGTGCGCCGT 5940  
QY 5941 GCAGCAATAGGAGCCTCAACCGATGTGAATGAATTAACAGATCTACAAAAGCATATGTA 6000  
DB 5941 GCAGCAATAGGAGCCTCAACCGATGTGAATGAATTAACAGATCTACAAAAGCATATGTA 6000  
QY 6001 AAAGATTTACAGATGTTGCTTAAAGAAAGAAACAGATGATTTATTTACTCAAGGCGAA 6060  
DB 6001 AAAGATTTACAGATGTTGCTTAAAGAAAGAAACAGATGATTTATTTACTCAAGGCGAA 6060  
QY 6061 GTAGTAAAGTGTAGATTAAGTATTCAAAAATCTTAATATTAAGAAAGCTTATACAA 6120  
DB 6061 GTAGTAAAGTGTAGATTAAGTATTCAAAAATCTTAATATTAAGAAAGCTTATACAA 6120  
QY 6121 AAAAGAAAAATAGTAATTAATAAAAGAAATTTGTTACCAATAGTCACTCATACTTTA 6180  
DB 6121 AAAAGAAAAATAGTAATTAATAAAAGAAATTTGTTACCAATAGTCACTCATACTTTA 6180  
QY 6181 AAATCTTATTTGGCAATTCGCGCTGTTCAGAGCAAGCCGAGTGGCAGAACTGTTAAT 6240  
DB 6181 AAATCTTATTTGGCAATTCGCGCTGTTCAGAGCAAGCCGAGTGGCAGAACTGTTAAT 6240  
QY 6241 ATCAACAAGGTTTATGAGAAACAGAAAGCTCTTGTAGAAAAATCTATTAATGCAAAA 6300  
DB 6241 ATCAACAAGGTTTATGAGAAACAGAAAGCTCTTGTAGAAAAATCTATTAATGCAAAA 6300  
QY 6301 CATTATTTCTGTAATATCAGAGATTAACAGAAATCAATCGAGTATAGAGTTCTGTTGAT 6360  
DB 6301 CATTATTTCTGTAATATCAGAGATTAACAGAAATCAATCGAGTATAGAGTTCTGTTGAT 6360  
QY 6361 GTTGTGGAATGTAAGAGTATGAGAGCTTCTTCTGATACCAATATTAATAAAGAAATACC 6420  
DB 6361 GTTGTGGAATGTAAGAGTATGAGAGCTTCTTCTGATACCAATATTAATAAAGAAATACC 6420  
QY 6421 AAGAACAAGGTTGAAAAAATCAATGTCGATGTAAGGTTTCGAGAAAGAGCTGAATTT 6480  
DB 6421 AAGAACAAGGTTGAAAAAATCAATGTCGATGTAAGGTTTCGAGAAAGAGCTGAATTT 6480  
QY 6481 ACAGCAGATTTCTAAGCAAGAAATTTCTTTTGAAGTCGAGATCGACAGCCGGGTA 6540  
DB 6481 ACAGCAGATTTCTAAGCAAGAAATTTCTTTTGAAGTCGAGATCGACAGCCGGGTA 6540  
QY 6541 GGAAGCCGAGAGTGCAGAAACCGTTTCCGTAAATCAATTTGCAAGAAAGACGGAATGAT 6600  
DB 6541 GGAAGCCGAGAGTGCAGAAACCGTTTCCGTAAATCAATTTGCAAGAAAGACGGAATGAT 6600  
QY 6601 GTGGAAGAAGCAAAAGATTTTGTATAAATAAAGCTGAGATTACAGCAAAAAGTTATGTTCT 6660  
DB 6601 GTGGAAGAAGCAAAAGATTTTGTATAAATAAAGCTGAGATTACAGCAAAAAGTTATGTTCT 6660  
QY 6661 GTTGCAATTTGAAATGCGCAGTGCAGAGTGCCTCAAAAGAGCTGGAATTTGAGACACA 6720  
DB 6661 GTTGCAATTTGAAATGCGCAGTGCAGAGTGCCTCAAAAGAGCTGGAATTTGAGACACA 6720  
QY 6721 GTGCGAGTTACCAAAAGATGAATCAACACGAGAGCAGAGTGAATAATTTCTTAAATTAAG 6780  
DB 6721 GTGCGAGTTACCAAAAGATGAATCAACACGAGAGCAGAGTGAATAATTTCTTAAATTAAG 6780  
QY 6781 ACTGGAACAAGATTAGATGTAATGACAGAAATGAGATTAATCAAGGTATCGGATCGGT 6840  
DB 6781 ACTGGAACAAGATTAGATGTAATGACAGAAATGAGATTAATCAAGGTATCGGATCGGT 6840  
QY 6841 TCAGCCGAGCTGGAATTTCTTGACGCCGAGTATCTGAGTGTGTTCTGTCAATTAATTT 6900

Db 6841 TCAGCCGAGAGCTGGAATTTCTTCAGCCGAGATCTGAGGTGTTCTCTCAATAATNT 6900  
Qy 6901 GCAGAAATAGAGTAAAGACAGATATCGATCACTTACCTTTCTATCGATGTAAAT 6960  
Db 6901 GCAGAAATAGAGTAAAGACAGATATCGATCACTTACCTTTCTATCGATGTAAAT 6960  
Qy 6961 GTAAAAAGCTCTTAATTAATTTTCGAAATTCCTTGAACAGCCGGTGGAGAGCCGAGTCTT 7020  
Db 6961 GTAAAAAGCTCTTAATTAATTTTCGAAATTCCTTGAACAGCCGGTGGAGAGCCGAGTCTT 7020  
Qy 7021 GGAGCAGTTACCGGAGTGGTTCTGTAAACATTAATAGTTCTGTAGTACCTGAGTT 7080  
Db 7021 GGAGCAGTTACCGGAGTGGTTCTGTAAACATTAATAGTTCTGTAGTACCTGAGTT 7080  
Qy 7081 CACAAATACCTGATTTGACTTCCGTACAGAAAAAGTAATGTAAACGCAAAAGAGNA 7140  
Db 7081 CACAAATACCTGATTTGACTTCCGTACAGAAAAAGTAATGTAAACGCAAAAGAGNA 7140  
Qy 7141 AAAAATATTAGCAAAACAGCAGCAATGCGAGATCGAGAGCAATCGAGCCAT 7200  
Db 7141 AAAAATATTAGCAAAACAGCAGCAATGCGAGATCGAGAGCAATCGAGCCAT 7200  
Qy 7201 GCTTGCTAATTAATTTTGGAAACGCTGTAGAAAGATAGAAAAATTCGAAGAAAAAGGA 7260  
Db 7201 GCTTGCTAATTAATTTTGGAAACGCTGTAGAAAGATAGAAAAATTCGAAGAAAAAGGA 7260  
Qy 7261 ACAGAAAGTTTAAAACTTTAGACGAAGTTAAACAAAGAACAGATAAAAATTAATCAT 7320  
Db 7261 ACAGAAAGTTTAAAACTTTAGACGAAGTTAAACAAAGAACAGATAAAAATTAATCAT 7320  
Qy 7321 GCTACGAAAAAAATCTTACAAATCAGCAGATTTCTACAGAAAGATCTTCTGTAAAAAGC 7380  
Db 7321 GCTACGAAAAAAATCTTACAAATCAGCAGATTTCTACAGAAAGATCTTCTGTAAAAAGC 7380  
Qy 7381 GATAGAGGATATCTCAGGAGAAAGAAATTAAGCCATTGTGAAGCTTCTGATNTAT 7440  
Db 7381 GATAGAGGATATCTCAGGAGAAAGAAATTAAGCCATTGTGAAGCTTCTGATNTAT 7440  
Qy 7441 GGAAGAAATGTAGATATTAACAACAGAGCAAGAAATATATCATCTTCTAGTGGTTTG 7500  
Db 7441 GGAAGAAATGTAGATATTAACAACAGAGCAAGAAATATATCATCTTCTAGTGGTTTG 7500  
Qy 7501 GGAAGCTCAGGCTCTGCTCCGATCAGGAAACGTGGCAGTTACAAATATTAAGAAAT 7560  
Db 7501 GGAAGCTCAGGCTCTGCTCCGATCAGGAAACGTGGCAGTTACAAATATTAAGAAAT 7560  
Qy 7561 TCCGAGGTTACTGTTGAAAAATCTTTGTGAAAAGCAGCTGMAAAAGTAATGTATGATCG 7620  
Db 7561 TCCGAGGTTACTGTTGAAAAATCTTTGTGAAAAGCAGCTGMAAAAGTAATGTATGATCG 7620  
Qy 7621 GATATTACAGGAAATGCTTTAAACAGCATCAAGGCTCTAGAGCATTTGGGAAT 7680  
Db 7621 GATATTACAGGAAATGCTTTAAACAGCATCAAGGCTCTAGAGCATTTGGGAAT 7680  
Qy 7681 GGAAGCTCAGTATGCAAAATTAATTTCTAATGGAAGATCAAAATATCAATTTAAATTTCT 7740  
Db 7681 GGAAGCTCAGTATGCAAAATTAATTTCTAATGGAAGATCAAAATATCAATTTAAATTTCT 7740  
Qy 7741 AAGCTATTAGAAAAAATATGATGTATTTAAAGATTAATCGAAATTTGAGACGGAA 7800  
Db 7741 AAGCTATTAGAAAAAATATGATGTATTTAAAGATTAATCGAAATTTGAGACGGAA 7800  
Qy 7801 GCAGAAAGATTAACGTAAGAGGCTAGCTCCGAGCCATTAATCTCAAAAGCAAGAAAT 7860  
Db 7801 GCAGAAAGATTAACGTAAGAGGCTAGCTCCGAGCCATTAATCTCAAAAGCAAGAAAT 7860  
Qy 7861 GAAATGAATTCAGAGGTTGAAATTTGAGAGAGTATTTTCAATGAGAAAAATAGATPACT 7920  
Db 7861 GAAATGAATTCAGAGGTTGAAATTTGAGAGAGTATTTTCAATGAGAAAAATAGATPACT 7920  
Qy 7921 AGCCCTTTCAAGGAATTTGAGAGAGAAATCATGTCAAAAGTGGAAAAAGAAACAGAGTG 7980

Db 7921 AGCCCTTTCAAGGAATTTGAGAGAGAAATCATGTCAAAAGTGGAAAAAGAAACAGAGTG 7980  
Qy 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGACAGAGGGGAGAAATTAATTCGGA 8040  
Db 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGACAGAGGGGAGAAATTAATTCGGA 8040  
Qy 8041 GCAAAAGATGCGGAAGCTCTTAATTTGAAAGTATGTAACAAATCCGGAAGAGTATTTT 8100  
Db 8041 GCAAAAGATGCGGAAGCTCTTAATTTGAAAGTATGTAACAAATCCGGAAGAGTATTTT 8100  
Qy 8101 CATGCAATTAATGTGAATATGAAAGCAACATTAATTAAGTAACAGAGTTCTAAA 8160  
Db 8101 CATGCAATTAATGTGAATATGAAAGCAACATTAATTAAGTAACAGAGTTCTAAA 8160  
Qy 8161 GCAGTAACAGGTTCTGTATTTGGAGAGATTGAGTACCAAGGCAAGCTTCTCTCA 8220  
Db 8161 GCAGTAACAGGTTCTGTATTTGGAGAGATTGAGTACCAAGGCAAGCTTCTCTCA 8220  
Qy 8221 GGTAAATCTATGATGAGAGTTGAGAGAGAAATTTGTCAGAACAAATCGATTGAATGCA 8280  
Db 8221 GGTAAATCTATGATGAGAGTTGAGAGAGAAATTTGTCAGAACAAATCGATTGAATGCA 8280  
Qy 8281 ATTTCTAAGTGAAGGTTTGGATGAGATTAAGTAACTGCTAAATCTTCTGTAGTATCA 8340  
Db 8281 ATTTCTAAGTGAAGGTTTGGATGAGATTAAGTAACTGCTAAATCTTCTGTAGTATCA 8340  
Qy 8341 GGAATTTGAGAGAGAAATTCGCGAGCAGAGTGAATCTTCTACGCAACAAAGTAACT 8400  
Db 8341 GGAATTTGAGAGAGAAATTCGCGAGCAGAGTGAATCTTCTACGCAACAAAGTAACT 8400  
Qy 8401 GAAATCCGTAAGTCCGTTTACGAAAGCAATTAAGAAATTAATGATTAACAAAAAATAT 8460  
Db 8401 GAAATCCGTAAGTCCGTTTACGAAAGCAATTAAGAAATTAATGATTAACAAAAAATAT 8460  
Qy 8461 ATTTCAGAAAGTCAATGCTCTGCTTAAATGATACAAAGAAAGACGATATAGATCT 8520  
Db 8461 ATTTCAGAAAGTCAATGCTCTGCTTAAATGATACAAAGAAAGACGATATAGATCT 8520  
Qy 8521 TTAAGCGTAGCCGATGTCATGACACAGAGCAACAAAGCAATTTACAGATCAACAG 8580  
Db 8521 TTAAGCGTAGCCGATGTCATGACACAGAGCAACAAAGCAATTTACAGATCAACAG 8580  
Qy 8581 TTAATCTTCAACACGTTAAATGAGAGAAACGATCTCAACTCGGCAAAACCTTGGCT 8640  
Db 8581 TTAATCTTCAACACGTTAAATGAGAGAAACGATCTCAACTCGGCAAAACCTTGGCT 8640  
Qy 8641 TTAATCTTCAACACGTTAAATGAGAGAAACGATCTCAACTCGGCAAAACCTTGGCT 8640  
Db 8641 TTAATCTTCAACACGTTAAATGAGAGAAACGATCTCAACTCGGCAAAACCTTGGCT 8640  
Qy 8641 AAAAATGAAAAATTAAGAAATGTAAGGAATGAGAGACCTTATGCGAGCGGAACA 8700  
Db 8641 AAAAATGAAAAATTAAGAAATGTAAGGAATGAGAGACCTTATGCGAGCGGAACA 8700  
Qy 8701 GCAGCCGTTGAAAAATTAACAAGAGTACTACAGAGCAATTTGTTGCAGAAATTTGGAA 8760  
Db 8701 GCAGCCGTTGAAAAATTAACAAGAGTACTACAGAGCAATTTGTTGCAGAAATTTGGAA 8760  
Qy 8761 ATTTGAGATTAATTTGAAACGATTTGCAAGAGTAAATGATTTGATCAAGAGAC 8820  
Db 8761 ATTTGAGATTAATTTGAAACGATTTGCAAGAGTAAATGATTTGATCAAGAGAC 8820  
Qy 8821 GGAACCAAGAGGCTTGTGGAAGAGATGTAATTTCTGTGAAAAATCAATTTCAAGG 8880  
Db 8821 GGAACCAAGAGGCTTGTGGAAGAGATGTAATTTCTGTGAAAAATCAATTTCAAGG 8880  
Qy 8881 GAAACAAATCATTCATTTGAATGAATAGCCAGAAATTTGGAACCGGAAGTAAATGTA 8940  
Db 8881 GAAACAAATCATTCATTTGAATGAATAGCCAGAAATTTGGAACCGGAAGTAAATGTA 8940  
Qy 8941 GATGCTTTGAATGAACCTTGAATGATCTACAGAGAAAAAGTGTGCTATGCTGAAT 9000  
Db 8941 GATGCTTTGAATGAACCTTGAATGATCTACAGAGAAAAAGTGTGCTATGCTGAAT 9000  
Qy 9001 GGTATTTGAAAAATGTAATGATATGATTTAAGAAAAATGTAAGCCAAAAATCGGA 9060  
Db 9001 GGTATTTGAAAAATGTAATGATATGATTTAAGAAAAATGTAAGCCAAAAATCGGA 9060

QY	9061	GGAGCATGCTATTTGTGAAAACCTACCTGAAAAACAAGAAATCAAGCATTTTACAGAGCAAAA	9122
Db	9061	AGACATGCTATTTGTGAAAACCTACCTGAAAAACAAGAAATCAAGCATTTTACAGAGCAAAA	9120
QY	9121	GTTAAATATCTTGTGAAAAAGAGACCGCTGACGCTGAGCTGCAATATCGAATGTACACTT	9180
Db	9121	GTTAAATATCTTGTGAAAAAGAGACCGCTGACGCTGAGCTGCAATATCGAATGTACACTT	9180
QY	9181	TCCAATGAGATGATATTAAAAATTTGGCAAGCAGTATGCATCTTCAATTAATAAC	9240
Db	9181	TCCAATGAGATGATATTAAAAATTTGGCAAGCAGTATGCATCTTCAATTAATAAC	9240
QY	9241	AAAAATTCAAAAAATAATTAATTAATCTTTAGATTCGAATGCAATGCAATGTAATGTTAT	9300
Db	9241	AAAAATTCAAAAAATAATTAATTAATCTTTAGATTCGAATGCAATGCAATGTAATGTTAT	9300
QY	9301	GCGGTGCGTGAAGCAAGAGGTGCGAGAGCCCAAGCGACAGCTTAGTGTAAAGATCAATA	9360
Db	9301	GCGGTGCGTGAAGCAAGAGGTGCGAGAGCCCAAGCGACAGCTTAGTGTAAAGATCAATA	9360
QY	9361	AATAGAACTAATATATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAAACTCAATGTA	9420
Db	9361	AATAGAACTAATATATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAAACTCAATGTA	9420
QY	9421	TATGCGGATATGATATAAAATTTATATATTAAGTAAAGAAATCTTAAGGCTATTGCGGAT	9480
Db	9421	TATGCGGATATGATATAAAATTTATATATTAAGTAAAGAAATCTTAAGGCTATTGCGGAT	9480
QY	9481	GCCAAAGTCATGCTGCGAGCTGCTTCGCGCACTGCGCACTATTGAAAAAATGAAGTAAAA	9540
Db	9481	GCCAAAGTCATGCTGCGAGCTGCTTCGCGCACTGCGCACTATTGAAAAAATGAAGTAAAA	9540
QY	9541	TTTAATTAATGCGATCCGAGAAATTTTAAAAATTAATCTGGCAAGTTGGAAAGGAAAGCTAAT	9600
Db	9541	TTTAATTAATGCGATCCGAGAAATTTTAAAAATTAATCTGGCAAGTTGGAAAGGAAAGCTAAT	9600
QY	9601	AAAAAAGTGGTGTGATCTTAATCAGGAGACTGGTATACGATTAATTAATTAATCATGCGAT	9660
Db	9601	AAAAAAGTGGTGTGATCTTAATCAGGAGACTGGTATACGATTAATTAATTAATCATGCGAT	9660
QY	9661	TCTTCTGAAAAAGCATACAAAAAATTTGACATATCATCAAGAGAGAGAAAAAGGAAAA	9720
Db	9661	TCTTCTGAAAAAGCATACAAAAAATTTGACATATCATCAAGAGAGAGAAAAAGGAAAA	9720
QY	9721	AAATGA 9726	
Db	9721	AAATGA 9726	

Query Match	100.0%;	Score 9726;	DB 3;	Length 11130;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 9726;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	TTGAGCGGCATCAAAAAATPAAGCTTCAGAGGCAAGAGGAAGAGATATTCAGATTCTTAATAAA	60
Db	1034	ATGAGCGGCATCAAAAAATPAAGCTTCAGAGGCAAGAGGAAGAGATATTCAGATTCTTAATAAA	1099
QY	61	GTTTATATGATTTTGGGATTGGTATTAACAATAAGCGTGAAGGCTTAATGATCAATC	120
Db	1094	GTTTATATGATTTTGGGATTGGTATTAACAATAAGCGTGAAGGCTTAATGATCAATC	1155
QY	121	ACCGGACTGAGAAATTTTGGAAACAAAAATAGAAAAAAGATATATGTTTATGACATTACT	180
Db	1154	ACCGGACTGAGAAATTTTGGAAACAAAAATAGAAAAAAGATATATGTTTATGACATTACT	1213
QY	181	ACAAACAAGATTCAAGGGGAGAAAGGCTTTTAAAGTTTATPAGATTGGCTTTAACAA	240
Db	1214	ACAAACAAGATTCAAGGGGAGAAAGGCTTTTAAAGTTTATPAGATTGGCTTTAACAA	1277
QY	241	AATATATAGCAAAATCTATATTTTGGGAAAAAGATAGTACGGGGGTAAATATCTTTT	300
Db	1274	AATATATAGCAAAATCTATATTTTGGGAAAAAGATAGTACGGGGGTAAATATCTTTT	1333
QY	301	AACTTTGTCAATGGAAAAATTTGAAGTAGATGGGATTATCAAACGGAATTCGAGAAATATA	360
Db	1334	AACTTTGTCAATGGAAAAATTTGAAGTAGATGGGATTATCAAACGGAATTCGAGAAATATA	1393
QY	361	ATTGAGAGAAATTTATATTTTCTTAAGCTCGAGAGGGATGGGAGTAGAGAAAAATTGAGATT	420
Db	1334	ATTGAGAGAAATTTATATTTTCTTAAGCTCGAGAGGGATGGGAGTAGAGAAAAATTGAGATT	1455
QY	421	ATCAATGCTGGTCTTTTCATTCTATATATTCGAAAAAAGATGATTTTAAAGAGCTTTG	480
Db	1454	ATCAATGCTGGTCTTTTCATTCTATATATTCGAAAAAAGATGATTTTAAAGAGCTTTG	1513
QY	481	GAAGAAGCCAAACATGTGTAAGTTTAAATGGAAATCATTCGACGTATGAAAAAGTAAAA	540
Db	1514	GAAGAAGCCAAACATGTGTAAGTTTAAATGGAAATCATTCGACGTATGAAAAAGTAAAA	1573
QY	541	ATTCCATTGATTCGGAATCGGAATGGAAGCATTCGCTAGAGAAAAATCAATGCTGTGGAAGC	600
Db	1574	ATTCCATTGATTCGGAATCGGAATGGAAGCATTCGCTAGAGAAAAATCAATGCTGTGGAAGC	1633
QY	601	ATCGGTTTATATGCGGCGGATATTAGATTGAAAGATACCTGCAATATCTAAGAACAAGAAAT	660
Db	1634	ATCGGTTTATATGCGGCGGATATTAGATTGAAAGATACCTGCAATATCTAAGAACAAGAAAT	1693
QY	661	ACAGATTTTAAAAATTTTATAGCAATATTTAGATGGAATAATTTCTGGCTGTACCGGAAAT	720
Db	1694	ACAGATTTTAAAAATTTTATAGCAATATTTAGATGGAATAATTTCTGGCTGTACCGGAAAT	1755
QY	721	TTAAAGGTCACCAAGACAAATCTGAGATATTATTTCTTCAGCTCACATAGATTCTCCT	780
Db	1754	TTAAAGGTCACCAAGACAAATCTGAGATATTATTTCTTCAGCTCACATAGATTCTCCT	1813
QY	781	CAAAAAGCTATGGGAAAAAATTCAACTGTTGAAAAGAGATAGAGAAATATGTATAAGGA	840
Db	1814	CAAAAAGCTATGGGAAAAAATTCAACTGTTGAAAAGAGATAGAGAAATATGTATAAGGA	1877
QY	841	AATACCAAGCAATATTTGGAATCTGATGCTGTATATTGAAAGCAATGGAATATRTAAAAAT	900
Db	1874	AATACCAAGCAATATTTGGAATCTGATGCTGTATATTGAAAGCAATGGAATATRTAAAAAT	1933
QY	901	AGTACGAAGCTACAAATGCGAGATTTTATAAGAAAGAGGGGAAAAAGAACTTATAAC	960
Db	1934	AGTACGAAGCTACAAATGCGAGATTTTATAAGAAAGAGGGGAAAAAGAACTTATAAC	1999
QY	961	ACTCCTTTAAGTTTATCAGATGTGGAAGCTTCGTAAGATTAATTAAGAAAAAGTCATA	1020
Db	1994	ACTCCTTTAAGTTTATCAGATGTGGAAGCTTCGTAAGATTAATTAAGAAAAAGTCATA	2053

QY	1021	GGAAAGATGTGGACATTTACAGCTGAAGCAAGAAATTTCTATGATCCAACTTTAGTTACT	1080
Db	2054	GGAAAGATGTGGACATTTACAGCTGAAGCAAGAAATTTCTATGATCCAACTTTAGTTACT	2113
QY	1081	AAGCTTCGAAAGCACTCTTTTACGTTTGTGTAACAGGTTCTAATTTCTCTATCCAACTTTAAAT	1140
Db	2114	AAGCTTCGAAAGCACTCTTTTACGTTTGTGTAACAGGTTCTAATTTCTCTATCCAACTTTAAAT	2173
QY	1141	GGATTTTATAGGTTATTTAGACAAAGTAAGTCCAGTGTGGTTATTTGAAAGAAAGTGCAAAGTC	1200
Db	2174	GGATTTTATAGGTTATTTAGACAAAGTAAGTCCAGTGTGGTTATTTGAAAGAAAGTGCAAAGTC	2233
QY	1201	GAAGCAACAGAGGAAAGGCMAATATTCATTCTTACAGTGGAGTAAGCACTATGGGA	1260
Db	2234	GAAGCAACAGAGGAAAGGCMAATATTCATTCTTACAGTGGAGTAAGCACTATGGGA	2293
QY	1261	GCACCTACTTCCATTTAAAAATTAACCAATTTATTTGAGAAAGCCAAATGGAAAACTT	1320
Db	2294	GCACCTACTTCCATTTAAAAATTAACCAATTTATTTGAGAAAGCCAAATGGAAAACTT	2353
QY	1321	CTCAGTATCGAGCGGGGATATATTTCTGCAAAAAAGTAATTCAAATGTAACTATTTGAAGA	1380
Db	2354	CTCAGTATCGAGCGGGGATATATTTCTGCAAAAAAGTAATTCAAATGTAACTATTTGAAGA	2413
QY	1381	GAAGTAAATCGAAGGGAAGAGCAGATATTCTTCAAAATCTGAAATTACTATTGATGCT	1440
Db	2414	GAAGTAAATCGAAGGGAAGAGCAGATATTCTTCAAAATCTGAAATTACTATTGATGCT	2473
QY	1441	TCTGTTTCTGTTGGAAAGATAGAGATTTCCAAATTAAGTACTCTTTCAAGTATTTGGAGAG	1500
Db	2474	TCTGTTTCTGTTGGAAAGATAGAGATTTCCAAATTAAGTACTCTTTCAAGTATTTGGAGAG	2533
QY	1501	GAAAGAGAAATTTAAATCTTCCGTCAAGTTGCTTAAAGAGCAAAAGTAGAATCAGAAACG	1560
Db	2534	GAAAGAGAAATTTAAATCTTCCGTCAAGTTGCTTAAAGAGCAAAAGTAGAATCAGAAACG	2593
QY	1561	GATGATGTAAATGTGAGAGAGTGAAGCGATTAATTTCAATTCGAGCTGCTGTAAAAAGTGGA	1620
Db	2594	GATGATGTAAATGTGAGAGAGTGAAGCGATTAATTTCAATTCGAGCTGCTGTAAAAAGTGGA	2653
QY	1621	TTGGGGGATAGTGGTAATGGGGTTGTGGCTGCGCAATTTTCTAATCTAATAGCTTCTCTCC	1680
Db	2654	TTGGGGGATAGTGGTAATGGGGTTGTGGCTGCGCAATTTTCTAATCTAATAGCTTCTCTCC	2713
QY	1681	CGTATAGATGTAGATGATATCTACATCTCCAAAGAGCGACTAAATGTGAGAGCTCATTAAC	1740
Db	2714	CGTATAGATGTAGATGATATCTACATCTCCAAAGAGCGACTAAATGTGAGAGCTCATTAAC	2773
QY	1741	ATTACTTAAAAATAGTGTTCTGCAACAGGATCTGATTTGGGAAGCTTCCAAAGTTATGAAT	1800
Db	2774	ATTACTTAAAAATAGTGTTCTGCAACAGGATCTGATTTGGGAAGCTTCCAAAGTTATGAAT	2833
QY	1801	GATCACGTTTATGAAATCAGGTCATCTTAAATCAATTTTATGATGCAATTTAAACAGCGGTTT	1860
Db	2834	GATCACGTTTATGAAATCAGGTCATCTTAAATCAATTTTATGATGCAATTTAAACAGCGGTTT	2893
QY	1861	GGAGGAGACAGTGCTAATGAGGAAATTAAGAATAAGCTAAGCACTTATTTAGTGTGGT	1920
Db	2894	GGAGGAGACAGTGCTAATGAGGAAATTAAGAATAAGCTAAGCACTTATTTAGTGTGGT	2953
QY	1921	GTGTCTGCAACCATAGCAAAATCATATATATTCTGCTTGTGGCAATGAGAGAGTGA	1980
Db	2954	GTGTCTGCAACCATAGCAAAATCATATATATTCTGCTTGTGGCAATGAGAGAGTGA	3013
QY	1981	AAGCTTTCTTCAAGAGTGGAAAGGAGTAAATGTAAGGGAATTAATGAAAGCTCAAAATCTT	2040
Db	3014	AAGCTTTCTTCAAGAGTGGAAAGGAGTAAATGTAAGGGAATTAATGAAAGCTCAAAATCTT	3073
QY	2041	CGAGCGACTACGCTCAAGTGAAGTGTGGCTGTGACAAAGAGAAAGAAAAAGAACTTATTT	2100
Db	3074	CGAGCGACTACGCTCAAGTGAAGTGTGGCTGTGACAAAGAGAAAGAAAAAGAACTTATTT	3133
QY	2101	GGAAATGACAGATTTTTTATGAGAACTATTAATAATTAATGCTTCTGTGCAATTTGCCAT	2160

Db	3134	GGAAATCACAGT	TTTTTATGAACT	AAAAATATCTCT	GTGACAA	TTGCCGAT	31193					
QY	2161	CATGCTGAA	TTGGTATCGA	AGAAAAAT	TGATATCA	ACGTGAAA	TTGAATAT	2220				
Db	3194	CATGCTGA	ATTGGTATCGA	AGAAAAAT	TGATATCA	ACGTGAAA	TTGAATAT	3253				
QY	2221	AAAAATCCTT	CAAAAAAT	TGCAAAAGT	CTGTAT	TATGAAT	TTAACA	CTTTTAA	GAGAGCT	2280		
Db	3254	AAAAATCCTT	CAAAAAAT	TGCAAAAGT	CTGTAT	TATGAAT	TTAACA	CTTTTAA	GAGAGCT	3113		
QY	2281	TTTGGAAAA	AGAAACGA	AAATCCG	AGATATGAT	CCGAAAGAT	AT	TGAACTAT	TGAAAAA	2340		
Db	3314	TTTGGAAAA	AGAAACGA	AAATCCG	AGATATGAT	CCGAAAGAT	AT	TGAACTAT	TGAAAAA	3373		
QY	2341	TTATTGAAT	GCATTTTC	CAGAAAAAT	TGATGAAAA	CCGAGCTTT	TACT	AAATG	GTGAA	2400		
Db	3374	TTATTGAAT	GCATTTTC	CAGAAAAAT	TGATGAAAA	CCGAGCTTT	TACT	AAATG	GTGAA	3433		
QY	2401	AGAAATG	CAAT	TATCTTCCG	AGTGAAC	TTTCAAAA	CAGAA	CTGCTAT	TGAAAT	TGCA	2460	
Db	3434	AGAAATG	CAAT	TATCTTCCG	AGTGAAC	TTTCAAAA	CAGAA	CTGCTAT	TGAAAT	TGCA	3493	
QY	2461	AACATATG	TCAGGAGAA	TGAAAAAT	TGAGAAAA	TTTACG	GAATTA	CCGAAAGAT	TTTAAAGCT	2520		
Db	3494	AACATATG	TCAGGAGAA	TGAAAAAT	TGAGAAAA	TTTACG	GAATTA	CCGAAAGAT	TTTAAAGCT	3553		
QY	2521	TTTTCAG	AGAT	TGAGTGA	CTGAT	TAAAGAA	CTTTGA	ATTTTAC	CAGAGT	TGAAAT	2580	
Db	3554	TTTTCAG	AGAT	TGAGTGA	CTGAT	TAAAGAA	CTTTGA	ATTTTAC	CAGAGT	TGAAAT	3613	
QY	2581	TATGCAAA	TTTTCACA	CTTTTAC	CTTCCG	AGCTAA	TGAGAA	AAAGAT	TGTTCT	CT	2640	
Db	3614	TATGCAAA	TTTTCACA	CTTTTAC	CTTCCG	AGCTAA	TGAGAA	AAAGAT	TGTTCT	CT	3673	
QY	2641	GTGGAGAG	AGCTGTTG	TGGGTGA	CAAGAA	ATTATAG	CAAGAT	TATCCG	TGAAAAA	2700		
Db	3674	GTGGAGAG	AGCTGTTG	TGGGTGA	CAAGAA	ATTATAG	CAAGAT	TATCCG	TGAAAAA	3733		
QY	2701	GGAGCTAAA	CTTGCTG	CAAAAAAG	ATTTAAAT	TAAAGCTAT	CAAT	TAAAG	CAGAA	CA	2766	
Db	3734	GGAGCTAAA	CTTGCTG	CAAAAAAG	ATTTAAAT	TAAAGCTAT	CAAT	TAAAG	CAGAA	CA	3793	
QY	2761	GTGAATTT	AGTTGGA	AATAT	TGGA	CTTGCG	AGAA	GCAGT	CAATCCG	GAAGTGC	AGTCCGA	2820
Db	3794	GTGAATTT	AGTTGGA	AATAT	TGGA	CTTGCG	AGAA	GCAGT	CAATCCG	GAAGTGC	AGTCCGA	3853
QY	2821	GGAAGTTAA	TATGTTCA	AAAGAT	TGCAAAAA	AT	CAGCTAT	CGTGAAG	AGCTTAA	GAAAAAGCT	2880	
Db	3854	GGAAGTTAA	TATGTTCA	AAAGAT	TGCAAAAA	AT	CAGCTAT	CGTGAAG	AGCTTAA	GAAAAAGCT	3913	
QY	2881	GAATTTAT	CAGAGAAA	AATAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	2940
Db	3914	GAATTTAT	CAGAGAAA	AATAT	TAT	TAT	TAT	TAT	TAT	TAT	TAT	3973
QY	2941	TCTTTTAA	TG	TG	TG	TG	TG	TG	TG	TG	TG	3000
Db	3974	TCTTTTAA	TG	TG	TG	TG	TG	TG	TG	TG	TG	4033
QY	3001	ATCAGTA	AGGCAAG	GTTC	CAAT	TGATG	ACGAA	GCAT	TTTG	GAAGCT	TAT	3060
Db	4034	ATCAGTA	AGGCAAG	GTTC	CAAT	TGATG	ACGAA	GCAT	TTTG	GAAGCT	TAT	4093
QY	3061	GCTTTTAA	CAGTAA	AGAT	TACT	TCTG	TGTA	TG	TG	TG	TG	3120
Db	4094	GCTTTTAA	CAGTAA	AGAT	TACT	TCTG	TGTA	TG	TG	TG	TG	4153
QY	3121	ACGAAAAA	TGCGGCG	TGCGG	TGCTG	TGCG	TGTA	TAT	TAT	TAT	TAT	3186
Db	4154	ACGAAAAA	TGCGGCG	TGCGG	TGCTG	TGCG	TGTA	TAT	TAT	TAT	TAT	4213
QY	3181	GCTTCAT	TGAAGAT	TATG	TGA	AGGCA	CAAA	GTAA	TAT	TAT	TAT	3246

Db 4214 GCTTCATTGAAGATTAATGACGAAGACAAGTAAATATGATTAAGATTAAGATGANA 4273  
Qy 3241 GTAACAGTAATCGCGGAATCTTTAGAGTAGATGCAAAAAACGACCGGAACAATCAACAGT 3300  
Db 4274 GTAACAGTAATCGCGGAATCTTTAGAGTAGATGCAAAAAACGACCGGAACAATCAACAGT 4333  
Qy 3301 ATTTCTGTTGCGGAGGAATTAATAAGTTGGAAGTAAACCGAGTGAAGAAAAACCGAA 3360  
Db 4334 ATTTCTGTTGCGGAGGAATTAATAAGTTGGAAGTAAACCGAGTGAAGAAAAACCGAA 4393  
Qy 3361 TCAGAGAAAGAACGACGAGGAGATTTTGTGCAAAATCGGAAACAAAGTGAAGCTGTGTA 3420  
Db 4394 TCAGAGAAAGAACGACGAGGAGATTTTGTGCAAAATCGGAAACAAAGTGAAGCTGTGTA 4453  
Qy 3421 AATAAATTTACGGATAGTATGATTCATTAACAGAAAAAATTACAAATTTACTGTGA 3480  
Db 4454 AATAAATTTACGGATAGTATGATTCATTAACAGAAAAAATTACAAATTTACTGTGA 4513  
Qy 3481 GGAAGTAAAAAAGCGGGGAATCTTCCTTCGAACGTTTCTCATATCTCCGATTAAGACCG 3540  
Db 4514 GGAAGTAAAAAAGCGGGGAATCTTCCTTCGAACGTTTCTCATATCTCCGATTAAGACCG 4573  
Qy 3541 TCTTTCAGTTTGGAGCTTCTGAAAGTGTCTTTCATATATTTAAAAAGAAACATCT 3600  
Db 4574 TCTTTCAGTTTGGAGCTTCTGAAAGTGTCTTTCATATATTTAAAAAGAAACATCT 4633  
Qy 3601 GCTGTCTGATGAGATTAAGATTAATTTGAAGGAGCAAAATTAAGGTAGAGTGAAGT 3660  
Db 4634 GCTGTCTGATGAGATTAAGATTAATTTGAAGGAGCAAAATTAAGGTAGAGTGAAGT 4693  
Qy 3661 TCTTCTGATTTACTTTTGTGAGACATGAGGCGGATCTGCTGCACTTCAGTGAATCAT 3720  
Db 4694 TCTTCTGATTTACTTTTGTGAGACATGAGGCGGATCTGCTGCACTTCAGTGAATCAT 4753  
Qy 3721 ATTGGAAGTGAATAATAGCAACATCAAGTGTGTTTATGCTGAGCGGCTCTGTAATAT 3780  
Db 4754 ATTGGAAGTGAATAATAGCAACATCAAGTGTGTTTATGCTGAGCGGCTCTGTAATAT 4813  
Qy 3781 ATTCAAGTGAATAACAAGTCTTGTGTAATAATAGTATTCGAATGCGCAATTAATTT 3840  
Db 4814 ATTCAAGTGAATAACAAGTCTTGTGTAATAATAGTATTCGAATGCGCAATTAATTT 4873  
Qy 3841 AAAAGTAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGCAGGTTTGAAGCAGT 3900  
Db 4874 AAAAGTAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGCAGGTTTGAAGCAGT 4933  
Qy 3901 AAAAGTAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGCAGGTTTGAAGCAGT 3960  
Db 4934 AAAAGTAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGCAGGTTTGAAGCAGT 4993  
Qy 3961 GTGAACAATGAAGTTTCTGCAAAATCAGAAAAATTAATACAGTAGCAGGAGAACTGAAAGC 4020  
Db 4994 GTGAACAATGAAGTTTCTGCAAAATCAGAAAAATTAATACAGTAGCAGGAGAACTGAAAGC 5053  
Qy 4021 CAAAAAATGATGTTGATGTCATCTGTTATCAAGCGGACACCCAAAGTGAACAGAGCTTTA 4080  
Db 5054 CAAAAAATGATGTTGATGTCATCTGTTATCAAGCGGACACCCAAAGTGAACAGAGCTTTA 5113  
Qy 4081 AATTTACAAAGCTGGAAGAGTCAATGGAATCTGTAAGGGGCTTACTGTGACTGTGCCAAATTA 4140  
Db 5114 AATTTACAAAGCTGGAAGAGTCAATGGAATCTGTAAGGGGCTTACTGTGACTGTGCCAAATTA 5173  
Qy 4141 AACACAAGTAATGCTTCTATAGTGTGGAAGATATACTAAGCTTAATGAGCGGAC 4200  
Db 5174 AACACAAGTAATGCTTCTATAGTGTGGAAGATATACTAAGCTTAATGAGCGGAC 5233  
Qy 4201 GCAAAAGCTCTTTAGCAACCACTCAAGTGAAGTGTGCTGCAAGTGAACGCGGAGCAAT 4260  
Db 5234 GCAAAAGCTCTTTAGCAACCACTCAAGTGAAGTGTGCTGCAAGTGAACGCGGAGCAAT 5293  
Qy 4261 AGTTCTGAGCGGAGTATGAAATTAATCAAGGGGCTGTTTCTGTCAATTAAGATTGACAT 4320  
Db 5294 AGTTCTGAGCGGAGTATGAAATTAATCAAGGGGCTGTTTCTGTCAATTAAGATTGACAT 5353

Qy 4321 GACGTGAAGCTAGCGTTGATTAATCTTCATCGAAGAGCTAAATGAATCAATGTCATT 4380  
Db 5354 GACGTGAAGCTAGCGTTGATTAATCTTCATCGAAGAGCTAAATGAATCAATGTCATT 5413  
Qy 4381 GCCAAGATGTCAAAGAGGTTCTGATCTACAAAAGAAATTCAGGCTTTACTAATGGA 4440  
Db 5414 GCCAAGATGTCAAAGAGGTTCTGATCTACAAAAGAAATTCAGGCTTTACTAATGGA 5473  
Qy 4441 AAAAGTAATAATTTTGAAGATCTGTGATTTAATTCAGTGAAGTGAATGTTAATGAG 4500  
Db 5474 AAAAGTAATAATTTTGAAGATCTGTGATTTAATTCAGTGAAGTGAATGTTAATGAG 5533  
Qy 4501 AAGAACACTAGAAAAAGCAAAAGAAAAAAGAGAGCGGCTCATTTGTAATGTCCTTTA 4560  
Db 5534 AAGAACACTAGAAAAAGCAAAAGAAAAAAGAGAGCGGCTCATTTGTAATGTCCTTTA 5593  
Qy 4561 TCGGTTGCTGGAACGGAATTAATCCGCTGAGAGAGTACTAATTCAGTCAATCTGTTAA 4620  
Db 5594 TCGGTTGCTGGAACGGAATTAATCCGCTGAGAGAGTACTAATTCAGTCAATCTGTTAA 5653  
Qy 4621 AATAAATTTAAGCAAGATTAAGTGAAGCAATTAAGAGCGGAGAGGATTAATATCAT 4680  
Db 5654 AATAAATTTAAGCAAGATTAAGTGAAGCAATTAAGAGCGGAGAGGATTAATATCAT 5713  
Qy 4681 GCGAAACATGTAATGTTGAGAGCAAAATCATCTAGTGTGTTGTAATGCGGCTTCTGGA 4740  
Db 5714 GCGAAACATGTAATGTTGAGAGCAAAATCATCTAGTGTGTTGTAATGCGGCTTCTGGA 5773  
Qy 4741 CTGCTATCAGCAAAAGATGCTTTTTCAGAAATGGAGTGGAGCATGGCAAGACTTATCA 4800  
Db 5774 CTGCTATCAGCAAAAGATGCTTTTTCAGAAATGGAGTGGAGCATGGCAAGACTTATCA 5833  
Qy 4801 AATGACACGATTTGCAAGAGTGGATTAAGGAAGATTTCTGTAATCTCTTAATGTAAC 4860  
Db 5834 AATGACACGATTTGCAAGAGTGGATTAAGGAAGATTTCTGTAATCTCTTAATGTAAC 5893  
Qy 4861 GCAAAATTAATTCATTTCTGGGGTGAATGTTCCGGGAACCATTCGCGGTTCTCTTCTACG 4920  
Db 5894 GCAAAATTAATTCATTTCTGGGGTGAATGTTCCGGGAACCATTCGCGGTTCTCTTCTACG 5953  
Qy 4921 GCGGTGAGAGCTGCTTTTTCGAATTAATCTCTTCAATAATAAACCTCTGCTTGAATCA 4980  
Db 5954 GCGGTGAGAGCTGCTTTTTCGAATTAATCTCTTCAATAATAAACCTCTGCTTGAATCA 6013  
Qy 4981 GGAACGAAGTAAATCCTTTAGTGAAGAAATCAAAAGTCAATGTACAAGCTTGAAT 5040  
Db 6014 GGAACGAAGTAAATCCTTTAGTGAAGAAATCAAAAGTCAATGTACAAGCTTGAAT 6073  
Qy 5041 GATTCCTATTTACAAACGTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGGCTGAATC 5100  
Db 6074 GATTCCTATTTACAAACGTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGGCTGAATC 6133  
Qy 5101 GAGAGATGATATCTGTCATATCGTGTTCGTATGAAACGGAAGCTTATAGTATGTTCT 5160  
Db 6134 GAGAGATGATATCTGTCATATCGTGTTCGTATGAAACGGAAGCTTATAGTATGTTCT 6193  
Qy 5161 GAGTTGAAGAGTAAGTCTTTTCAATGTAATGCAAAAGATTCAAAATTAATCA 5220  
Db 6194 GAGTTGAAGAGTAAGTCTTTTCAATGTAATGCAAAAGATTCAAAATTAATCA 6253  
Qy 5221 ATTGCGGAAATGCAAAATGAGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTATGCA 5280  
Db 6254 ATTGCGGAAATGCAAAATGAGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTATGCA 6313  
Qy 5281 AATATTTGAAAAACAATCATGTTATAGCTATTTGTAATAAAAAACAGTAATAATTAACA 5340  
Db 6314 AATATTTGAAAAACAATCATGTTATAGCTATTTGTAATAAAAAACAGTAATAATTAACA 6373  
Qy 5341 GATCAAGTGAAGAAAAATATCAATGTAATGCAAAAGTTATATCTATGACCAATATCTATA 5400  
Db 6374 GATCAAGTGAAGAAAAATATCAATGTAATGCAAAAGTTATATCTATGACCAATATCTATA 6433

QY	5401	GCAGTCCGAGTTGGAGAGCAAAAGAGCCTCTGTGCAGAGAGCTTCTGCAGTACTACC	5460
Db	6434	GCAGTCGGAGTTGGAGAGCAAAAAGAGCCTCTGTGCAGAGAGCTTCTGCAGTACTACC	6493
QY	5461	TTGAAATAGACAGTTCTCTCTCATGTTGATCCAACTGATATTGACAAGATTTAGAGAA	5520
Db	6494	TTGAAATAGACAGTTCTCTCTCATGTTGATCCAACTGATATTGACAAGATTTAGAGAA	6553
QY	5521	GAATAATATGAAATTAAGAAAAGCAAAATGTTATGTTAGCTGAAAAATCAGAGTCAA	5580
Db	6554	GAATAATATGAAATTAAGAAAAGCAAAATGTTATGTTAGCTGAAAAATCAGAGTCAA	6613
QY	5561	GTGTCACAAATGCGACACGTCTTCCGAGCAAGTGGACAAGTGCAGTAGAGAGCTGGA	5640
Db	6614	GTGTCACAAATGCGACACGTCTTCCGAGCAAGTGGACAAGTGCAGTAGAGAGCTGGA	6673
QY	5611	GTACAGCTTAATTAATTAACAACAATATCTCTGCACTATTAATAAATTAATTAATTAAT	5700
Db	6674	GTACAGCTTAATTAATTAACAACAATATCTCTGCACTATTAATAAATTAATTAATTAAT	6733
QY	5701	GTACGAAATGCTTTGGTTAAAGCAAAATCTCATTCATCTATTAACAATTTGGAATTGGA	5760
Db	6734	GTACGAAATGCTTTGGTTAAAGCAAAATCTCATTCATCTATTAACAATTTGGAATTGGA	6793
QY	5761	GCTGGAGTTGAGCTGGAGAGCTGGAGAGTGCAGAGTTCGTAGCAGTAATTAAGATTGTA	5820
Db	6794	GCTGGAGTTGAGCTGGAGAGCTGGAGAGTGCAGAGTTCGTAGCAGTAATTAAGATTGTA	6853
QY	5821	AATAATACGATATGACGAATTTAATATCAGCAAAATTCATCTGGAGAGGAATATGTCGGAGTT	5880
Db	6854	AATAATACGATATGACGAATTTAATATCAGCAAAATTCATCTGGAGAGGAATATGTCGGAGTT	6913
QY	5881	ATTACAGAGTCTGATGCGGTAAATGCTAATTATATCAGGAACAGTGTGAGAGTGCCTGT	5940
Db	6914	ATTACAGAGTCTGATGCGGTAAATGCTAATTATATCAGGAACAGTGTGAGAGTGCCTGT	6973
QY	5941	GCAGCAATAGAGAGCCTCAACAGTGTGATGAATTAACAGGATCTACAAAACATATGTGA	6000
Db	6974	GCAGCAATAGAGAGCCTCAACAGTGTGATGAATTAACAGGATCTACAAAACATATGTGA	7033
QY	6001	AAAGATTTCTACAGTATGTTGCTTAAGAGAGAAACAGATGATTTATTAATCTACAGGGCAA	6060
Db	7034	AAAGATTTCTACAGTATGTTGCTTAAGAGAGAAACAGATGATTTATTAATCTACAGGGCAA	7093
QY	6061	GTAGATTAAGTGTGATTAAGTATTTCAAAATCTTAATATTAACGAAGACTTATCACAA	6120
Db	7094	GTAGATTAAGTGTGATTAAGTATTTCAAAATCTTAATATTAACGAAGACTTATCACAA	7153
QY	6121	AAAAAATAATAGTATTAATTAAGAAATGTTGTACCAATAGTCACTCATATCTTTA	6180
Db	7154	AAAAAATAATAGTATTAATTAAGAAATGTTGTACCAATAGTCACTCATATCTTTA	7213
QY	6181	AAATCTTTATTTGGCAATGCGCGTGTTCAGACCAAGCCGAGTGGCAGAGAACTGTTAAT	6240
Db	7214	AAATCTTTATTTGGCAATGCGCGTGTTCAGACCAAGCCGAGTGGCAGAGAACTGTTAAT	7273
QY	6241	ATCAACAAGGTTTATGGAAGAAACAGAGCTCTGTGAGAAATTTCTAATTATTAATGCAAAA	6300
Db	7274	ATCAACAAGGTTTATGGAAGAAACAGAGCTCTGTGAGAAATTTCTAATTATTAATGCAAAA	7333
QY	6301	CATTATCTGTAAATTCAGAGATTTACAGAAATTCMAATCGAGATAGTAGTCTGTGGT	6360
Db	7334	CATTATCTGTAAATTCAGAGATTTACAGAAATTCMAATCGAGATAGTAGTCTGTGGT	7393
QY	6361	GTTTGTGAAATGTTAGAGATAGAGACTTCTCTGATATCCAAATATTATTAAGAAATATCC	6420
Db	7394	GTTTGTGAAATGTTAGAGATAGAGACTTCTCTGATATCCAAATATTATTAAGAAATATCC	7453
QY	6421	AAGACAAGATGGAAGAAATCTACATGTCGTATGAAGTTTCCGAGAGAAAGCTGAATTT	6480
Db	7454	AAGACAAGATGGAAGAAATCTACATGTCGTATGAAGTTTCCGAGAGAAAGCTGAATTT	7513
QY	6481	ACAGCAGATTCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTCCGACAGCCGGGGTA	6540

Db	7514	ACACGAGATTCTAAGCAGGAAATTTCTCTTTTGGAGTCGGAATCCGACGACCGGGGTA	7573
Qy	6541	GGAGCCGGAGTGGCAGGAACCGTTCCGTAAATCAATTTGCAGAAAGACGAAATGAT	6600
Db	7574	GGAGCCGGAATGGCAGGAACCGTTCCGTAAATCAATTTGCAGAAAGACGAAATGAT	7633
Qy	6601	GTGGAGAGCAAAAGATTTTGGTAAAAAAGCTGAGATTACACCAAAACCTTATAGTTC	6660
Db	7634	GTGGAGAGCAAAAGATTTTGGTAAAAAAGCTGAGATTACACCAAAACCTTATAGTTC	7693
Qy	6661	GTTCGCAATTGGAATGCGCAGTCGAGTGGCTGCAGAAAAGAGCTGGAAATTCGACGCA	6720
Db	7694	GTTCGCAATTGGAATGCGCAGTCGAGTGGCTGCAGAAAAGAGCTGGAAATTCGACGCA	7753
Qy	6721	GTGCGCAATTACCAAAAGATGAAATCAACGAGACGAAGATGAAAAATTTCTAAATTTATG	6780
Db	7754	GTGCGCAATTACCAAAAGATGAAATCAACGAGACGAAGATGAAAAATTTCTAAATTTATG	7813
Qy	6781	ACTCGAAACAAAGTTAGATGTAAATGCGAAATAGAGATAAATTCAGTACTCGAATCGCT	6840
Db	7814	ACTCGAAACAAAGTTAGATGTAAATGCGAAATAGAGATAAATTCAGTACTCGAATCGCT	7873
Qy	6841	TCAGCCGAGCTGGAAATCTTGACGCGCGAGTATCTGAGTGGTTCTGTCAATTAATTT	6900
Db	7874	TCAGCCGAGCTGGAAATCTTGACGCGCGAGTATCTGAGTGGTTCTGTCAATTAATTT	7933
Qy	6901	GCAATATAGGTAGAAACAGATATCGATCATAGTACTTTACCTTCTACTGATGTAAAT	6960
Db	7934	GCAATATAGGTAGAAACAGATATCGATCATAGTACTTTACCTTCTACTGATGTAAAT	7993
Qy	6961	GTAAGACCTCTTAATAAATTTTGCATTTCTTGACAGCCGCTGGAGAGCCCGACGCTTT	7020
Db	7994	GTAAGACCTCTCTTAATAAATTTTGCATTTCTTGACAGCCGCTGGAGAGCCCGACGCTTT	8053
Qy	7021	GCACACATTCACCGAGTGGTTCTGTTAAACATTAATAGTTCTGTGATAGCTCGAGTT	7080
Db	8054	GCACACATTCACCGAGTGGTTCTGTTAAACATTAATAGTTCTGTGATAGCTCGAGTT	8113
Qy	7081	CACATATACTCTGATTTGACTTCCGTACGAGAAAAGATAATGTTACGGCAAAAGAGAA	7140
Db	8114	CACATATACTCTGATTTGACTTCCGTACGAGAAAAGATAATGTTACGGCAAAAGAGAA	8173
Qy	7141	AAAAATATTAAGCAACAGCAGCAAAATGCGAAATCGAGAGACGCAATCGAGCCCAAT	7200
Db	8174	AAAAATATTAAGCAACAGCAGCAAAATGCGAAATCGAGAGACGCAATCGAGCCCAAT	8233
Qy	7201	GTCTTGGTAAATATTTTGGAAACGCTGTGAAGATGAAAAATTTCTGAAAGGAAAAAGGA	7260
Db	8234	GTCTTGGTAAATATTTTGGAAACGCTGTGAAGATGAAAAATTTCTGAAAGGAAAAAGGA	8293
Qy	7261	ACAGAGTTTAAAAAACTTTAGACGAGTTTAAACAAACACAGATTAATAAAGTAAATGAT	7320
Db	8294	ACAGAGTTTAAAAAACTTTAGACGAGTTTAAACAAACACAGATTAATAAAGTAAATGAT	8353
Qy	7321	GCTACGAAAAAAATCTTACATTCAGCAGGATTTTCTACAGAGATCTTCTGTAAACCG	7380
Db	8354	GCTACGAAAAAAATCTTACATTCAGCAGGATTTTCTACAGAGATCTTCTGTAAACCG	8413
Qy	7381	GATAGAGAGATCTCAGGAGGAAGAAATTAACCATTTGGAGACTTCTGATATTAATTT	7440
Db	8414	GATAGAGAGATCTCAGGAGGAAGAAATTAACCATTTGGAGACTTCTGATATTAATTT	8473
Qy	7441	GGAAAAAATGTAGATATTACACAGAGACCAAGATATAATCACCCTTCTACTGCTGTTTG	7500
Db	8474	GGAAAAAATGTAGATATTACACAGAGACCAAGATATAATCACCCTTCTACTGCTGTTTG	8533
Qy	7501	GGAACTGAGGCTCTTGCTTCCGCATTCAGGAACAGTGGCAGTTTACAAATATTTAAAGAAAT	7560
Db	8534	GGAACTGAGGCTCTTGCTTCCGCATTCAGGAACAGTGGCAGTTTACAAATATTTAAAGAAAT	8593
Qy	7561	TCCGAGTTACTGTTGAAATTTCTTTTGTGAACAGCTGAAAAAGTAAATGTTAGATCG	7620

Db 8594 TCCGAGTTACTGTTGAAAAATCTTTTGTGAAAGCACTGAAAAAGTAAATTTTGATCG 8653  
Qy 7621 GATATTACAGAAATGTTGCTTTTAAACAGCATATCAAGGCTCTGTAGAGCATTTGGAAATA 7680  
Db 8654 GATATTACAGAAATGTTGCTTTTAAACAGCATATCAAGGCTCTGTAGAGCATTTGGAAATA 8713  
Qy 7681 GAGAGTGCCTATGACAGAAATTAATTTCTAATGGAAGATCAAAATATCAATTTTAAATTCCT 7740  
Db 8714 GAGAGTGCCTATGACAGAAATTAATTTCTAATGGAAGATCAAAATATCAATTTTAAATTCCT 8773  
Qy 7741 AAGCTATTGAAAAAAATATGATGTTTGTAAAAGATAAATCCGAATTTGAGCGGAA 7800  
Db 8774 AAGCTATTGAAAAAAATATGATGTTTGTAAAAGATAAATCCGAATTTGAGCGGAA 8833  
Qy 7801 GCAAAAAGATTAAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAGAT 7860  
Db 8834 GCAAAAAGATTAAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAGAT 8893  
Qy 7861 GAAATGAAATTCAGAGTTGAAATTTGAGAGAGATTTTTCATGAAAGAAATAGAGTAACT 7920  
Db 8894 GAAATGAAATTCAGAGTTGAAATTTGAGAGAGATTTTTCATGAAAGAAATAGAGTAACT 8953  
Qy 7921 AGCCCTTTCAAGGAATTGAAAGAGAAATCAATGTCAAAGTGAAAAAGAAACAGAGTG 7980  
Db 8954 AGCCCTTTCAAGGAATTGAAAGAGAAATCAATGTCAAAGTGAAAAAGAAACAGAGTG 9013  
Qy 7981 ACTGCTGAATCTCAAGAGACTTCTGTAGAGCAGTAGCAGGGCAGAAATTTATTTCCGA 8040  
Db 9014 ACTGCTGAATCTCAAGAGACTTCTGTAGAGCAGTAGCAGGGCAGAAATTTATTTCCGA 9073  
Qy 8041 GCAAAAAGATGCGGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100  
Db 9074 GCAAAAAGATGCGGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 9133  
Qy 8101 CATGCAGATATGATGAAATTTGAGAGCAACATTAATGAAATGAAAGTAAACAGCAGTTCTAA 8160  
Db 9134 CATGCAGATATGATGAAATTTGAGAGCAACATTAATGAAATGAAAGTAAACAGCAGTTCTAA 9193  
Qy 8161 GCAGTAAACAGGTTCTGTATTTGGAGAGTTGGAGTCAACCAAGCAGAACTTCTGCTCA 8220  
Db 9194 GCAGTAAACAGGTTCTGTATTTGGAGAGTTGGAGTCAACCAAGCAGAACTTCTGCTCA 9253  
Qy 8221 GGTAAAACTATGATGAGAGTTGAGAGAGAAATTTGTTCAAGCAAAATCGATTGAATGCA 8280  
Db 9254 GGTAAAACTATGATGAGAGTTGAGAGAGAAATTTGTTCAAGCAAAATCGATTGAATGCA 9313  
Qy 8281 ATTTCTAAAGTAGAAGTTGATGAGAAATAAGTAACTGTAATCTTCTGTAGATCA 8340  
Db 9314 ATTTCTAAAGTAGAAGTTGATGAGAAATAAGTAACTGTAATCTTCTGTAGATCA 9373  
Qy 8341 GGAATGAGAGAGATTGCGCGAGCAGAGTGAATCTTCTTACAGCACAAAGTAACT 8400  
Db 9374 GGAATGAGAGAGATTGCGCGAGCAGAGTGAATCTTCTTACAGCACAAAGTAACT 9433  
Qy 8401 GAAATCCGTAAGTTGTTTACGAAAGCAAGTTATGAAATATGATTTACCAAAAAATAT 8460  
Db 9434 GAAATCCGTAAGTTGTTTACGAAAGCAAGTTATGAAATATGATTTACCAAAAAATAT 9493  
Qy 8461 ATTTCAGAGAGTAAATGCTTGTCTTAAATGATCAAAAGATGAAGCGGAAATTAATCT 8520  
Db 9494 ATTTCAGAGAGTAAATGCTTGTCTTAAATGATCAAAAGATGAAGCGGAAATTAATCT 9553  
Qy 8521 TTAGCGGATGCGCGATGTGATGCACAAAGAGCAAAAGCAATTTTACAGATCAAAAG 8580  
Db 9554 TTAGCGGATGCGCGATGTGATGCACAAAGAGCAAAAGCAATTTTACAGATCAAAAG 9613  
Qy 8581 TTAACTTTTACAACTGTAAATGAGAGAAAGTATCTCAACTTGTGCAAAAGCTTTGCT 8640  
Db 9614 TTAACTTTTACAACTGTAAATGAGAGAAAGTATCTCAACTTGTGCAAAAGCTTTGCT 9673  
Qy 8641 AAAAAATGAAAAATTAAGAGAAATGTAAGAGAACTGAGAGAGCTTGTGAGAGAGAGAAACA 8700  
Db 9674 AAAAAATGAAAAATTAAGAGAAATGTAAGAGAACTGAGAGAGCTTGTGAGAGAGAGAAACA 9733

Qy 8701 GCAGCGTTGAAAAATTAATACAAAGTACTACAGAGCATTTGTTGACAGAAATTTGGAA 8760  
Db 9734 GCAGCGTTGAAAAATTAATACAAAGTACTACAGAGCATTTGTTGACAGAAATTTGGAA 9793  
Qy 8761 ATTGAGATTAATTAAGAAAGCATTTGCAAGAGATTAATAGATTTGTAAGTCAACGAGAAC 8820  
Db 9794 ATTGAGATTAATTAAGAAAGCATTTGCAAGAGATTAATAGATTTGTAAGTCAACGAGAAC 9853  
Qy 8821 GGAACCAAGAGAGTCTGTGCGAAAGAGATGTAATTTCTGTGAAAAATATCAATTTCAAGG 8880  
Db 9854 GGAACCAAGAGAGTCTGTGCGAAAGAGATGTAATTTCTGTGAAAAATATCAATTTCAAGG 9913  
Qy 8881 GAAACAAATATCATTCATTGAAGATTAAGCCAGAAATTTGTGAACCGGAAGTGAATGTA 8940  
Db 9914 GAAACAAATATCATTCATTGAAGATTAAGCCAGAAATTTGTGAACCGGAAGTGAATGTA 9973  
Qy 8941 GATGCTTTGAATGAACCTTGATGATGTTTACAAAGAAAAATGCTGCTATGCTGGAATTT 9000  
Db 9974 GATGCTTTGAATGAACCTTGATGATGTTTACAAAGAAAAATGCTGCTATGCTGGAATTT 10033  
Qy 9001 GGTATTGGAATTTGATGATGTAATTAATGATTAAGAAAAATGTAAGAACCAAAATCGGA 9060  
Db 10034 GGTATTGGAATTTGATGATGTAATTAATGATTAAGAAAAATGTAAGAACCAAAATCGGA 10093  
Qy 9061 AGACATGCTATTGTAGAAATCTAGAAAAACAAGATTAACAAGCATTTTACAAGACAAAA 9120  
Db 10094 AGACATGCTATTGTAGAAATCTAGAAAAACAAGATTAACAAGCATTTTACAAGACAAAA 10153  
Qy 9121 GTAATATTTCTTGAAAAAGAGACGCTGCACTGACGTGCAATATGCAATGTCACATT 9180  
Db 10154 GTAATATTTCTTGAAAAAGAGACGCTGCACTGACGTGCAATATGCAATGTCACATT 10213  
Qy 9181 TCCATGAGATGATTAATTAATTTTGGCAAGCAGTATGATCTTCTCAATTAATACC 9240  
Db 10214 TCCATGAGATGATTAATTAATTTTGGCAAGCAGTATGATCTTCTCAATTAATACC 10273  
Qy 9241 AAAAAATTCAAAAATTAATATTACTTTAGCATCAAGTATGAAATGCAATGTTTCAT 9300  
Db 10274 AAAAAATTCAAAAATTAATATTACTTTAGCATCAAGTATGAAATGCAATGTTTCAT 10333  
Qy 9301 GGGGTGCTGAAGCAAGAGGTGCGCAGAGCCAAAGCAGTATGTTAAAGATCAAAATA 9360  
Db 10334 GGGGTGCTGAAGCAAGAGGTGCGCAGAGCCAAAGCAGTATGTTAAAGATCAAAATA 10393  
Qy 9361 AATGAACTAATTAATGTTGATTTTACAGAGAAAAATTAACAGAGGAAACATCAATGTA 9420  
Db 10394 AATGAACTAATTAATGTTGATTTTACAGAGAAAAATTAACAGAGGAAACATCAATGTA 10453  
Qy 9421 TATGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480  
Db 10454 TATGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10513  
Qy 9481 GCCAAAGTCAATGCTGCGACGCTGCTGCGCACTGCGCATATTGAAAAAAATGAAGTAAA 9540  
Db 10514 GCCAAAGTCAATGCTGCGACGCTGCTGCGCACTGCGCATATTGAAAAAAATGAAGTAAA 10573  
Qy 9541 TTTAATTAATGAGATCCGGAATTTAAAAATATATGCGCAAGATTTGAAAGGAAAGCTAAT 9600  
Db 10574 TTTAATTAATGAGATCCGGAATTTAAAAATATATGCGCAAGATTTGAAAGGAAAGCTAAT 10633  
Qy 9601 AAAAAACGTGCGTATGATCTTAATCAAGTATGATGTAACGATTAATTAATTAATTAAT 9660  
Db 10634 AAAAAACGTGCGTATGATCTTAATCAAGTATGATGTAACGATTAATTAATTAATTAAT 10693  
Qy 9661 TCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGAGAA 9720  
Db 10694 TCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGAGAA 10753  
Qy 9721 AAATGA 9726  
Db 10754 AAATGA 10759

RESULT 3  
 US-09-841-786-10  
 ; Sequence 10, Application US/09941786  
 ; Patent No. 6669940  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAGARAJA, T. G.  
 ; APPLICANT: STEWART, GEORGE C.  
 ; APPLICANT: NARAYANAN, SANDEEV K.  
 ; APPLICANT: CHENGAPPA, M. M.  
 ; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
 ; FILE REFERENCE: VACCINE AND PREPARATION THEREOF  
 ; CURRENT APPLICATION NUMBER: US/09/841,786  
 ; CURRENT FILING DATE: 2001-04-24  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 2780  
 ; TYPE: DNA  
 ; ORGANISM: Fusobacterium necrophorum  
 ; US-09-841-786-10

Query Match 28.6%; Score 2780; DB 3; Length 2780;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 GGGAGATTATTAAGAAAGGGGAAAAAAGAACTTATTAACCTCTTAAAGTTATCA 978  
 |||||  
 DB 1 GGGAGATTATTAAGAAAGGGGAAAAAAGAACTTATTAACCTCTTAAAGTTATCA 60  
 QY 979 GATGTGGAAGCTTCCGTAAAGATTAATAAGAAAAAGTCAATAGAAAGATGTTCAT 1038  
 |||||  
 DB 61 GATGTGGAAGCTTCCGTAAAGATTAATAAGAAAAAGTCAATAGAAAGATGTTCAT 120  
 QY 1039 ACAGCTGAAGCAAAAGATTTCTATGATGCACTTTAGTACTAGCTTGCAAGCACTCT 1098  
 |||||  
 DB 121 ACAGCTGAAGCAAAAGATTTCTATGATGCACTTTAGTACTAGCTTGCAAGCACTCT 180  
 QY 1099 TTTAGCTTTGTTACAGGTTCTATTTCTCCATCAATTTAATGGATTTTAAAGTTAT 1158  
 |||||  
 DB 181 TTTAGCTTTGTTACAGGTTCTATTTCTCCATCAATTTAATGGATTTTAAAGTTAT 240  
 QY 1159 ACAAGTAAGTCAGTGTCTGTTATTTGAAAAAGATGCCAAAGTCAGAACCAAGAGAAAG 1218  
 |||||  
 DB 241 ACAAGTAAGTCAGTGTCTGTTATTTGAAAAAGATGCCAAAGTCAGAACCAAGAGAAAG 300  
 QY 1219 GCAAAATATTCATTTCTTACAGTGAAGTAAAGCACTATGGAGACACTTCTCATTA 1278  
 |||||  
 DB 301 GCAAAATATTCATTTCTTACAGTGAAGTAAAGCACTATGGAGACACTTCTCATTA 360  
 QY 1279 AAAAATACCAATTTATTTTGGAGAAAGCCAAATGSAAACTTCTCAGTATCGGAGGGGA 1338  
 |||||  
 DB 361 AAAAATACCAATTTATTTTGGAGAAAGCCAAATGSAAACTTCTCAGTATCGGAGGGGA 420  
 QY 1339 TATATTTCTGCAAAAAGTAAATTCCTAATGTAACTATTGAAGAGAAATTAATCGAAGGA 1398  
 |||||  
 DB 421 TATATTTCTGCAAAAAGTAAATTCCTAATGTAACTATTGAAGAGAAATTAATCGAAGGA 480  
 QY 1399 AGAGCAGATATTACTTCAAAATCTGAAATATCTATTGATGCTTCTGTTCTGTTGGAACG 1458  
 |||||  
 DB 481 AGAGCAGATATTACTTCAAAATCTGAAATATCTATTGATGCTTCTGTTCTGTTGGAACG 540  
 QY 1459 ATAGAGATTTCCAAATTAAGAGCTTTTCAATGTTAGTGAACGAAAGAAATTAATCT 1518  
 |||||  
 DB 541 ATAGAGATTTCCAAATTAAGAGCTTTTCAATGTTAGTGAACGAAAGAAATTAATCT 600  
 QY 1519 TCCGTCAAGATTGCTAAAGAGCAAAAGTAAATCGAATCAAGAACGATGATGTAATGTGAGA 1578  
 |||||  
 DB 601 TCCGTCAAGATTGCTAAAGAGCAAAAGTAAATCGAATCAAGAACGATGATGTAATGTGAGA 660

QY 1579 AGTGAAGGATTAATTCATTCGAGCTGCTGTAAGAGGATGATGGGATAGTGTAT 1638  
 |||||  
 DB 661 AGTGAAGGATTAATTCATTCGAGCTGCTGTAAGAGGATGATGGGATAGTGTAT 720  
 QY 1639 GGGGTTGTGGCTGCAAAATATTTCTAATTAATGCTTCTCCGTATAGATAGATGGA 1698  
 |||||  
 DB 721 GGGGTTGTGGCTGCAAAATATTTCTAATTAATGCTTCTCCGTATAGATAGATGGA 780  
 QY 1699 TATCTACATGCCAAGACGACTAAATGTGAGGCTCATTAATCACTAAAAATGCTGT 1758  
 |||||  
 DB 781 TATCTACATGCCAAGACGACTAAATGTGAGGCTCATTAATCACTAAAAATGCTGT 840  
 QY 1759 CTGCAAAAGCATCTGATTTGGGAACCTTCAAGTTATGAAATGATCAGTTATGAATCA 1818  
 |||||  
 DB 841 CTGCAAAAGCATCTGATTTGGGAACCTTCAAGTTATGAAATGATCAGTTATGAATCA 900  
 QY 1819 GGTCAATCTAAATCAATTTTATGATGCAATTAACAAGCGGTTTGAAGACAGTGTCAAT 1878  
 |||||  
 DB 901 GGTCAATCTAAATCAATTTTATGATGCAATTAACAAGCGGTTTGAAGACAGTGTCAAT 960  
 QY 1879 GAGGAATTAAGAAATTAAGCTAAGCACTTATTTAGTGTGCTGTGCTCCCAACCAATAGA 1938  
 |||||  
 DB 961 GAGGAATTAAGAAATTAAGCTAAGCACTTATTTAGTGTGCTGTGCTCCCAACCAATAGA 1020  
 QY 1939 AATCAATTAATTTCTGCTTGTGGCAATGAGAGAGATGGAAGACTTCTTCAAGAGTG 1998  
 |||||  
 DB 1021 AATCAATTAATTTCTGCTTGTGGCAATGAGAGAGATGGAAGACTTCTTCAAGAGTG 1080  
 QY 1999 GAAGGAGATTAATGTAAGGCAATTAATGAGCTCAAAATCTTGAAGCACTACGTCACAT 2058  
 |||||  
 DB 1081 GAAGGAGATTAATGTAAGGCAATTAATGAGCTCAAAATCTTGAAGCACTACGTCACAT 1140  
 QY 2059 GGAAGTGTGCTGTACGAAGAGAGAAAAAACTTATTTGAATGCAAGCACTTTT 2118  
 |||||  
 DB 1141 GGAAGTGTGCTGTACGAAGAGAGAAAAAACTTATTTGAATGCAAGCACTTTT 1200  
 QY 2119 TATGAAATCTAATAAATTAATGCTTCTGTAACAATTCGATCATGCTGAATTTGATCG 2178  
 |||||  
 DB 1201 TATGAAATCTAATAAATTAATGCTTCTGTAACAATTCGATCATGCTGAATTTGATCG 1260  
 QY 2179 GAAGGAAAAATTTGATATCAACAGTGAATAAATTTGAATATTAATCTTCAAAATG 2238  
 |||||  
 DB 1261 GAAGGAAAAATTTGATATCAACAGTGAATAAATTTGAATATTAATCTTCAAAATG 1320  
 QY 2239 GCAAAAGCTGTTATTTATTAATTAATTAACCTTTTAAAGAGCTTTTGAAGAAAGCAAA 2298  
 |||||  
 DB 1321 GCAAAAGCTGTTATTTATTAATTAATTAACCTTTTAAAGAGCTTTTGAAGAAAGCAAA 1380  
 QY 2299 ACTCCAGAAATAGATCCGAAAGATTAATGAACTATTTGAAGAAATTAATGAAATGCAATTTCA 2358  
 |||||  
 DB 1381 ACTCCAGAAATAGATCCGAAAGATTAATGAACTATTTGAAGAAATTAATGAAATGCAATTTCA 1440  
 QY 2359 GAAAAATTTGATGAGAAAAACCGAGCTTTTAACTAAATGTTGAAGAAATGCAATTAATCTT 2418  
 |||||  
 DB 1441 GAAAAATTTGATGAGAAAAACCGAGCTTTTAACTAAATGTTGAAGAAATGCAATTAATCTT 1500  
 QY 2419 CCGGATGGAATCTTCAAAAACAGAACTGCTATAGAAATTTGCAACTATGTTCAAGGAGAA 2478  
 |||||  
 DB 1501 CCGGATGGAATCTTCAAAAACAGAACTGCTATAGAAATTTGCAACTATGTTCAAGGAGAA 1560  
 QY 2479 ATGAAAAATTTAGAGAAAAATTTACGAAAGAAATTTAAAGCTTTTCAAGAAATTTGAGT 2538  
 |||||  
 DB 1561 ATGAAAAATTTAGAGAAAAATTTACGAAAGAAATTTAAAGCTTTTCAAGAAATTTGAGT 1620  
 QY 2539 GGAATGATTAAGAAATTTTGAATTTTACAGAGATAGAAATTTATGCAATTTTCAACT 2598  
 |||||  
 DB 1621 GGAATGATTAAGAAATTTTGAATTTTACAGAGATAGAAATTTATGCAATTTTCAACT 1680  
 QY 2599 TTTACTCTTCTCGAGCTAATGAGAAAGATGTTTCTTCTGTGGAGAGAGCTGTTTGG 2658  
 |||||  
 DB 1681 TTTACTCTTCTCGAGCTAATGAGAAAGATGTTTCTTCTGTGGAGAGAGCTGTTTGG 1740  
 QY 2659 TGGGTAGAACAGAGAAATTTATGCAAGTATCCGTTGAAAAAGAGCTAAACTGTGTGCA 2718

Db 1741 TGGGTGAGACAGAGAAATTAAGCAAGGATCGTTGGAAAAAGAGCTAACTTGCTGCA 1800  
QY 2719 AAAAAAGATTAAATTAATAAGCTATCAATAAGCAAGAAACGTAAATTTATTTGGAAT 2778  
Db 1801 AAAAAAGATTAAATTAATAAGCTATCAATAAGCAAGAAACGTAAATTTATTTGGAAT 1860  
QY 2779 ATTGACCTTGCAGAAAGCAGTACATCCGAGAGTGCAGTGCAGAAAGATTAAATGTTCA 2838  
Db 1861 ATTGACCTTGCAGAAAGCAGTACATCCGAGAGTGCAGTGCAGAAAGATTAAATGTTCA 1920  
QY 2839 AGATGAAAAATTCACTATCTGTAGAGCTAAAGAAAAAGCTGAATTATCAGAGAAAAAT 2898  
Db 1921 AGATGAAAAATTCACTATCTGTAGAGCTAAAGAAAAAGCTGAATTATCAGAGAAAAAT 1980  
QY 2899 ATTAATGACAGATGACATTGACAGCTTTTCAATGAGCGGGATCTTTTAATGTTGCTCA 2958  
Db 1981 ATTAATGACAGATGACATTGACAGCTTTTCAATGAGCGGGATCTTTTAATGTTGCTCA 2040  
QY 2959 GGTGGGAAATGCAATCAATGAGATGGGAAGTTATAGTGAAGTATCAGTAAGGCAAGATT 3018  
Db 2041 GGTGGGAAATGCAATCAATGAGATGGGAAGTTATAGTGAAGTATCAGTAAGGCAAGATT 2100  
QY 3019 TCCATTGATGACGAAGCATTTTGAAGCTAAATAAAAAATTCCTTTAAACAGTAAGAT 3078  
Db 2101 TCCATTGATGACGAAGCATTTTGAAGCTAAATAAAAAATTCCTTTAAACAGTAAGAT 2160  
QY 3079 GATACCTTGCTTGGAAATGCTGCGGTTGACGGGAAATCGGAAACGAAAAATCGCGCGTC 3138  
Db 2161 GATACCTTGCTTGGAAATGCTGCGGTTGACGGGAAATCGGAAACGAAAAATCGCGCGTC 2220  
QY 3139 GGGGTTGCTGTTGCGGTAAATGATTAATGATTTTCAAAACAAAGCTTCCATTGAAGTAAT 3198  
Db 2221 GGGGTTGCTGTTGCGGTAAATGATTAATGATTTTCAAAACAAAGCTTCCATTGAAGTAAT 2280  
QY 3199 GACGAAGACAAAGTAATATGATTAAGATTAAGATGATGAAAGTAACAGTAACCTCGGAA 3258  
Db 2281 GACGAAGACAAAGTAATATGATTAAGATTAAGATGATGAAAGTAACAGTAACCTCGGAA 2340  
QY 3259 TCTTTAAGAGTGAAGCAAAAAACGACCGGAAACATCAACAGTAATTCGTGTCGCGAGGA 3318  
Db 2341 TCTTTAAGAGTGAAGCAAAAAACGACCGGAAACATCAACAGTAATTCGTGTCGCGAGGA 2400  
QY 3319 ATTAATAAGTGAAGTAACCGAGTGAAGAAAAACGAAATCAGAAAGAAAGACAGAG 3378  
Db 2401 ATTAATAAGTGAAGTAACCGAGTGAAGAAAAACGAAATCAGAAAGAAAGACAGAG 2460  
QY 3379 GGAATTTTGGCAAAATCGGAAACAAAGTGAAGCTCTGTAATAATTAATTCGAGTAAT 3438  
Db 2461 GGAATTTTGGCAAAATCGGAAACAAAGTGAAGCTCTGTAATAATTAATTCGAGTAAT 2520  
QY 3439 ATGATTCATTAACGAAAAAATTAACAATTACATTTTCGAGAGGTAAAAAACCGGGG 3498  
Db 2521 ATGATTCATTAACGAAAAAATTAACAATTACATTTTCGAGAGGTAAAAAACCGGGG 2580  
QY 3499 AATCTTCCCTCGAAGCTTTCATCTCCGATTAAGGACCGTCTTTCAGTTGGAGACT 3558  
Db 2581 AATCTTCCCTCGAAGCTTTCATCTCCGATTAAGGACCGTCTTTCAGTTGGAGACT 2640  
QY 3559 TCTGGAAGTGTCTTTCATTAATAATTAATAAGAAAAACATCTGCTGCTAGATGAGTA 3618  
Db 2641 TCTGGAAGTGTCTTTCATTAATAATTAATAAGAAAAACATCTGCTGCTAGATGAGTA 2700  
QY 3619 AAGATAAATTTGAAGGAGCAATAAAAAAGTGAAGGTGACCTTCTTGAATCTCTCTTT 3678  
Db 2701 AAGATAAATTTGAAGGAGCAATAAAAAAGTGAAGGTGACCTTCTTGAATCTCTCTTT 2760  
QY 3679 GTTGAAGCATGGGCGGATC 3698  
Db 2761 GTTGAAGCATGGGCGGATC 2780

RESULT 4

US-09-841-786-13  
; Sequence 13, Application US/09841786  
; Patent No. 6669940  
; GENERAL INFORMATION:  
; APPLICANT: NAGARAJA, T. G.  
; APPLICANT: STEWART, GEORGE C.  
; APPLICANT: NARAYANAN, SANJEEV K.  
; APPLICANT: CHENGAPPA, M. M.  
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF  
; FILE REFERENCE: 30296  
; CURRENT APPLICATION NUMBER: US/09/841,786  
; CURRENT FILING DATE: 2001-04-24  
; PRIOR APPLICATION NUMBER: 09/558,257  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 2322  
; TYPE: DNA  
; ORGANISM: Fusobacterium necrophorum  
US-09-841-786-13

Query Match 23.9%; Score 2322; DB 3; Length 2322;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7405 GGAATTAAGCATTTGGAACACTTCTGATATTAATGGAATAAAATGATGATTTACACA 7464  
Db 1 GGAATTAAGCATTTGGAACACTTCTGATATTAATGGAATAAAATGATGATTTACACA 60  
QY 7465 GAGACAAGATTAATATCACTTCTACCTGCTTGGAACTGCAAGCTCTTGCTCCGA 7524  
Db 61 GAGACAAGATTAATATCACTTCTACCTGCTTGGAACTGCAAGCTCTTGCTCCGA 120  
QY 7525 TCAGGACAGTGGAGTTACAAATATTAAGAAATTCGAGATTACTGTTGAAATTC 7584  
Db 121 TCAGGACAGTGGAGTTACAAATATTAAGAAATTCGAGATTACTGTTGAAATTC 180  
QY 7585 TTTGTGAAAGAGCTGGAATAATGTTGATTCGATATTAACGAAATGTTGCTTTA 7644  
Db 181 TTTGTGAAAGAGCTGGAATAATGTTGATTCGATATTAACGAAATGTTGCTTTA 240  
QY 7645 ACAGCATATCAAGGCTCTGTAGAGCATTTGGGAATAGAGCTGCTATGCAGAATTAAT 7704  
Db 241 ACAGCATATCAAGGCTCTGTAGAGCATTTGGGAATAGAGCTGCTATGCAGAATTAAT 300  
QY 7705 TCTAATGGAAGATCAAAATATCAGTAATTAATAATTCAGCTAATTAAGAAAAATATGAT 7764  
Db 301 TCTAATGGAAGATCAAAATATCAGTAATTAATAATTCAGCTAATTAAGAAAAATATGAT 360  
QY 7765 GTTATTTGTAAGATTAATTCGAAATTTGAGACCGGAAGCAATTAACCGTGAAGAG 7824  
Db 361 GTTATTTGTAAGATTAATTCGAAATTTGAGACCGGAAGCAATTAACCGTGAAGAG 420  
QY 7825 GTAGCTGCCGAGCATTTATCTCAAAAGCAAGAATGAATTCAGAGGTGAATTT 7884  
Db 421 GTAGCTGCCGAGCATTTATCTCAAAAGCAAGAATGAATTCAGAGGTGAATTT 480  
QY 7885 GAGAAGATTAATTTCAATGAAGAAATAGAGTAATCAGCCCTTCTAAAGAAATGGAAGA 7944  
Db 481 GAGAAGATTAATTTCAATGAAGAAATAGAGTAATCAGCCCTTCTAAAGAAATGGAAGA 540  
QY 7945 GAAATCAATGTCAGAGTGAAGAAAAAGAAACAGAGTGAATTCAGAGGCTTCT 8004  
Db 541 GAAATCAATGTCAGAGTGAAGAAAAAGAAACAGAGTGAATTCAGAGGCTTCT 600  
QY 8005 GTAGAGCAGTAGAGAGGAGCAGAAATTTTCGAGCAAGCAAAAGATCCGGAAGCTCTTAT 8064  
Db 601 GTAGAGCAGTAGAGAGGAGCAGAAATTTTCGAGCAAGCAAAAGATCCGGAAGCTCTTAT 660  
QY 8065 TTGAAAGTTAGTACAAATCCGGAAGAGTATTTTTCATGCAATGATGATTAATGGA 8124

Db 661 TTGAAGTATGATCAAAATCCGGAAGAGATATTTTTCATGCAATATATGGA 720  
 Qy 8125 GCAACATTAATAAGTAAGTAACAGAGTTCTTAAGACATACAGGTTCTGTATGGGA 8184  
 Db 721 GCAACATTAATAAGTAAGTAACAGAGTTCTTAAGACATACAGGTTCTGTATGGGA 780  
 Qy 8185 GGAGTTGAGTCAACCAAGGACGAGTCTGCTGAGGTTAAACATATGTTGAAGTTAG 8244  
 Db 781 GGAGTTGAGTCAACCAAGGACGAGTCTGCTGAGGTTAAACATATGTTGAAGTTAG 840  
 Qy 8245 GAAGGAATTTGTTGAGAACAAATGATGATGCAATTTCTAAGTGAAGGTTGGAT 8304  
 Db 841 GAAGGAATTTGTTGAGAACAAATGATGATGCAATTTCTAAGTGAAGGTTGGAT 900  
 Qy 8305 GAAGTAAGTAACTGCTAAATCTTCTGTATGATGAGAAATGAGAGAAATGCGGA 8364  
 Db 901 GAAGTAAGTAACTGCTAAATCTTCTGTATGATGAGAAATGAGAGAAATGCGGA 960  
 Qy 8365 GCAGAGTGAATATCTTCTACAGACAAAGTAATCTGAAATCCGTATGTTAGGAAG 8424  
 Db 961 GCAGAGTGAATATCTTCTACAGACAAAGTAATCTGAAATCCGTATGTTAGGAAG 1020  
 Qy 8425 CAAGATTAAGAAATATGATTAACAATAATTTTGAAGTCAATGCTCTGCT 8484  
 Db 1021 CAAGATTAAGAAATATGATTAACAATAATTTTGAAGTCAATGCTCTGCT 1080  
 Qy 8485 TTAATATGATTAACAATAAGTAAGTAAGTAAGTCTTGAAGCCGTATGATGCA 8544  
 Db 1081 TTAATATGATTAACAATAAGTAAGTAAGTAAGTCTTGAAGCCGTATGATGCA 1140  
 Qy 8545 CAAGGAACAACAAGCATTTTGAAGTAAACAAGTATCTTCAACTCACTGAATGA 8604  
 Db 1141 CAAGGAACAACAAGCATTTTGAAGTAAACAAGTATCTTCAACTCACTGAATGA 1200  
 Qy 8605 GGAAACGTATCTCAACTGTCGCAAAAGCTTGGCTTAAATAATGAATGAATGA 8664  
 Db 1201 GGAAACGTATCTCAACTGTCGCAAAAGCTTGGCTTAAATAATGAATGAATGA 1260  
 Qy 8665 AAAGAACTGAGAGAGCCCTTAACTCGAGCGGAACAGACGCGCTTGAATAAAG 8724  
 Db 1261 AAAGAACTGAGAGAGCCCTTAACTCGAGCGGAACAGACGCGCTTGAATAAAG 1320  
 Qy 8725 AGTACTACAGAGCATTTGTTGACAGAAATTTGGGAATTTGAGATTAATAAGCAT 8784  
 Db 1321 AGTACTACAGAGCATTTGTTGACAGAAATTTGGGAATTTGAGATTAATAAGCAT 1380  
 Qy 8785 GGAAGATTAATACGATTTGTAAGTCAACGAGACGGAACCAAGAGAGCTTGTGGA 8844  
 Db 1381 GGAAGATTAATACGATTTGTAAGTCAACGAGACGGAACCAAGAGAGCTTGTGGA 1440  
 Qy 8845 AAGATGATTTCTGTGAAAAATCAATTTCAAGGGAACAATAATCATCATTTGAAGAT 8904  
 Db 1441 AAGATGATTTCTGTGAAAAATCAATTTCAAGGGAACAATAATCATCATTTGAAGAT 1500  
 Qy 8905 AAAGCCAGAAATTTGGAACCGGAAGTGAATGTAAGTCTTGAAGTGAAGTGA 8964  
 Db 1501 AAAGCCAGAAATTTGGAACCGGAAGTGAATGTAAGTCTTGAAGTGAAGTGA 1560  
 Qy 8965 GATCTTCAAGGAAGAAAGTGTGCTATGTTGGAATGTTGGAATGTTGATGAAT 9024  
 Db 1561 GATCTTCAAGGAAGAAAGTGTGCTATGTTGGAATGTTGGAATGTTGATGAAT 1620  
 Qy 9025 AATGATTAAGAAAAATGTAAGCAAAATCGAAGACATGCTATTTAGAACTACT 9084  
 Db 1621 AATGATTAAGAAAAATGTAAGCAAAATCGAAGACATGCTATTTAGAACTACT 1680  
 Qy 9085 GGAAGAACAAAGATATACAGCATTTACAGAGCAAAAGTAATTTTGGAAAAAGAGAC 9144  
 Db 1681 GGAAGAACAAAGATATACAGCATTTACAGAGCAAAAGTAATTTTGGAAAAAGAGAC 1740  
 Qy 9145 GGTGAGCTGAGCTGATATGGAATGTAACATTTCAATGAGATGATTAATAAAT 9204  
 Db 1741 GGTGAGCTGAGCTGATATGGAATGTAACATTTCAATGAGATGATTAATAAAT 1800

Qy 9205 TTGGCAAGAGATATGATCTTCTCAATTAATACCAAAATTCAAAAATTAATATTACT 9264  
 Db 1801 TTGGCAAGAGATATGATCTTCTCAATTAATACCAAAATTCAAAAATTAATATTACT 1860  
 Qy 9265 TTAGCATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9324  
 Db 1861 TTAGCATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
 Qy 9325 GGAGCCAAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9384  
 Db 1921 GGAGCCAAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
 Qy 9385 GCAGAAAAAATTAACAGAGGGAACATCAATGATATGCGGATGATGATGATGAT 9444  
 Db 1981 GCAGAAAAAATTAACAGAGGGAACATCAATGATATGCGGATGATGATGATGAT 2040  
 Qy 9445 AATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9504  
 Db 2041 AATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
 Qy 9505 TCGCAACTGCACTATTTGAAAAAATGAAGTAAATTTAATGCGATCCGAAATTT 9564  
 Db 2101 TCGCAACTGCACTATTTGAAAAAATGAAGTAAATTTAATGCGATCCGAAATTT 2160  
 Qy 9565 AAAAATATCTGCAAGATTTGGAAGGAAAGCTAATTAATAAGTGGTGAATCTAAT 9624  
 Db 2161 AAAAATATCTGCAAGATTTGGAAGGAAAGCTAATTAATAAGTGGTGAATCTAAT 2220  
 Qy 9625 CAGTAGACTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9684  
 Db 2221 CAGTAGACTGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
 Qy 9685 TTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9726  
 Db 2281 TTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2322

RESULT 5  
 US-09-841-786-11  
 ; Sequence 11, Application US/09841786  
 ; Patent No. 6669940  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAGARAJA, T. G.  
 ; APPLICANT: STEWART, GEORGE C.  
 ; APPLICANT: NARAYANAN, SANJEEV K.  
 ; APPLICANT: CHENGABPA, M. M.  
 ; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
 ; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF  
 ; FILE REFERENCE: 30296  
 ; CURRENT APPLICATION NUMBER: US/09/841,786  
 ; CURRENT FILING DATE: 2001-04-24  
 ; PRIOR APPLICATION NUMBER: 09/558,257  
 ; PRIOR FILING DATE: 2000-04-25  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 11  
 ; LENGTH: 2141  
 ; TYPE: DNA  
 ; ORGANISM: Fusobacterium necrophorum  
 US-09-841-786-11

Query Match 22.0%; Score 2141; DB 3; Length 2141;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2141; Conservative 0; Indels 0; Gaps 0;

Qy 3553 GGAGCTTCTGGAAGGTTCTTCTCAATTAATTAATAAGAAACATCTGCTGCTAGAT 3612  
 Db 1 GGAGCTTCTGGAAGGTTCTTCTCAATTAATTAATAAGAAACATCTGCTGCTAGAT 60  
 Qy 3613 GGAGTAAGATTAATTTGAAGGAGCAATTAATAAGTGAAGTGAATCTTCTGATCT 3672  
 Db 61 GGAGTAAGATTAATTTGAAGGAGCAATTAATAAGTGAAGTGAATCTTCTGATCT 120

QY 3673 ACTTTGTTGAGCATGGGGCGGATCTGTCGACTTCAGTGGAAATCATATTGGAAGTGA 3732  
 Db 121 ACTTTGTTGAGCATGGGGCGGATCTGTCGACTTCAGTGGAAATCATATTGGAAGTGA 180  
 QY 3733 AATTAGACATCAAGCTGCTGTTAGTGGAGCGGCTGCTGTAATTAATTCAAGTAA 3792  
 Db 181 AATTAGACATCAAGCTGCTGTTAGTGGAGCGGCTGCTGTAATTAATTCAAGTAA 240  
 QY 3793 ACAAGTCTTGGTTAAAAATAGTGAATATTGAAAAATGCCAATAAATTTAAATGAT 3852  
 Db 241 ACAAGTCTTGGTTAAAAATAGTGAATATTGAAAAATGCCAATAAATTTAAATGAT 300  
 QY 3853 TTGAGTGAAGAACTCAAGTACAGCAGAGCAGGTTTGAAGCACTTAAAGAAAGTGA 3912  
 Db 301 TTGAGTGAAGAACTCAAGTACAGCAGAGCAGGTTTGAAGCACTTAAAGAAAGTGA 360  
 QY 3913 GGAACAAGAAAAAGTTATCTATTGGGAACCTTGCTCTGATCACTAGTGAACATGA 3972  
 Db 361 GGAACAAGAAAAAGTTATCTATTGGGAACCTTGCTCTGATCACTAGTGAACATGA 420  
 QY 3973 GTTTCGCAAAATCGAAAAATAACAGTACAGAGAGAACTGAAAGCCAAAAATGAT 4032  
 Db 421 GTTTCGCAAAATCGAAAAATAACAGTACAGAGAGAACTGAAAGCCAAAAATGAT 480  
 QY 4033 GTTGAATGTCAGCTTATCAAGCGGACACCAAGTGAACAGAGCTTAAATTACAGCT 4092  
 Db 481 GTTGAATGTCAGCTTATCAAGCGGACACCAAGTGAACAGAGCTTAAATTACAGCT 540  
 QY 4093 GGAAGATCAATGGAACCTGTAAGGCTACTGTAAGCTGTTGCCAATTAACAACAAGTA 4152  
 Db 541 GGAAGATCAATGGAACCTGTAAGGCTACTGTAAGCTGTTGCCAATTAACAACAAGTA 600  
 QY 4153 AATGCTTCTATTAGTGGGAGATTAATAAGTTAATCGAGCGGACCAAGCTCTT 4212  
 Db 601 AATGCTTCTATTAGTGGGAGATTAATAAGTTAATCGAGCGGACCAAGCTCTT 660  
 QY 4213 TTAGCAACCACTCAAGTGACTGTCAGTGAACGAGCGGAGCAATTAGTCTGAGAG 4272  
 Db 661 TTAGCAACCACTCAAGTGACTGTCAGTGAACGAGCGGAGCAATTAGTCTGAGAG 720  
 QY 4273 GGAATTAAGAAATTAACAAGGCTGTTTCTGTCATTAAGATTGAACATGACCTGAAGCT 4332  
 Db 721 GGAATTAAGAAATTAACAAGGCTGTTTCTGTCATTAAGATTGAACATGACCTGAAGCT 780  
 QY 4333 AGCGTTGATTAATCTTCATCGAAGGAGCTAATGAATGATGTCACCAAGATGTC 4392  
 Db 781 AGCGTTGATTAATCTTCATCGAAGGAGCTAATGAATGATGTCACCAAGATGTC 840  
 QY 4393 AAAAGAAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGAAGAAATGAAAA 4452  
 Db 841 AAAAGAAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGAAGAAATGAAAA 900  
 QY 4453 TATTTAAGAAAGATCGTGGATTTAATACGACTGGAATAGTTATTATACGAAGAACACTA 4512  
 Db 901 TATTTAAGAAAGATCGTGGATTTAATACGACTGGAATAGTTATTATACGAAGAACACTA 960  
 QY 4513 GAAAAAGCAAAAGAAAAAGAGAGCGGCTGTTAAATGCTGTTAATCGGTTGCTGGA 4572  
 Db 961 GAAAAAGCAAAAGAAAAAGAGAGCGGCTGTTAAATGCTGTTAATCGGTTGCTGGA 1020  
 QY 4573 ACGGATTAATCCGCTGGAAGAGTATGATTAAGCAATGCTGTTAAATTAATTTAA 4632  
 Db 1021 ACGGATTAATCCGCTGGAAGAGTATGATTAAGCAATGCTGTTAAATTAATTTAA 1080  
 QY 4633 GCAAGATTAGTGAAGCAATTAAGAACCGGAGAGAGATTAATTCATGCGAACATGTA 4692  
 Db 1081 GCAAGATTAGTGAAGCAATTAAGAACCGGAGAGAGATTAATTCATGCGAACATGTA 1140  
 QY 4693 AATGTTGAGGCAAAATCATCTACTGTTGTTGGAATCGGCTTCTGGAATGCTGTCAGC 4752  
 Db 1141 AATGTTGAGGCAAAATCATCTACTGTTGTTGGAATCGGCTTCTGGAATGCTGTCAGC 1200

QY 4753 AAAGATGCTTTTTCAGGAATGGATCTGAGACATGCAAGAATTATCAATGACACGATT 4812  
 Db 1201 AAAGATGCTTTTTCAGGAATGGATCTGAGACATGCAAGAATTATCAATGACACGATT 1260  
 QY 4813 GCAAGGTGATTAAGGAAGAAATTTCTGCTGATTTCTTAAATGGAACGCAAAATTTCC 4872  
 Db 1261 GCAAGGTGATTAAGGAAGAAATTTCTGCTGATTTCTTAAATGGAACGCAAAATTTCC 1320  
 QY 4873 ATTCTGGGAGTAAATGTTGCGGGAACATTGCGCGGTTCTCTTTTACGCGCGGTAGAGCT 4932  
 Db 1321 ATTCTGGGAGTAAATGTTGCGGGAACATTGCGCGGTTCTCTTTTACGCGCGGTAGAGCT 1380  
 QY 4933 GCTTTTGGAATTAATCTTTCATTAATAAAACCTCTGCTTTGATTACAGGAACGAAGTA 4992  
 Db 1381 GCTTTTGGAATTAATCTTTCATTAATAAAACCTCTGCTTTGATTACAGGAACGAAGTA 1440  
 QY 4993 AATCCTTTTATGGAAGAAATTAACAAAGTCAATGTAACAAGTTGAAATGATTTCTATTT 5052  
 Db 1441 AATCCTTTTATGGAAGAAATTAACAAAGTCAATGTAACAAGTTGAAATGATTTCTATTT 1500  
 QY 5053 ACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGAATCGAGGAATGTTA 5112  
 Db 1501 ACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGAATCGAGGAATGTTA 1560  
 QY 5113 TCTGTCATCGTGGTTCTGATGAACCGAAGCTTTAGTTAGTGAATCTGAGTTGAAAGA 5172  
 Db 1561 TCTGTCATCGTGGTTCTGATGAACCGAAGCTTTAGTTAGTGAATCTGAGTTGAAAGA 1620  
 QY 5173 GTTAAGTTCTTCAATGTAATGATGCAAAAGATTAACAAATTAATCAATTTGCCGGAAT 5232  
 Db 1621 GTTAAGTTCTTCAATGTAATGATGCAAAAGATTAACAAATTAATCAATTTGCCGGAAT 1680  
 QY 5233 GCAATGGAAGAAAAAGCGCTGAGGTTGGAACCAAGTGTCTCAATCAATTAATGAAAA 5292  
 Db 1681 GCAATGGAAGAAAAAGCGCTGAGGTTGGAACCAAGTGTCTCAATCAATTAATGAAAA 1740  
 QY 5293 CAATCAGTTAATGCTATTGTAATAAACAAGTAATTAACAAGCGGAATGATCAAGATGA 5352  
 Db 1741 CAATCAGTTAATGCTATTGTAATAAACAAGTAATTAACAAGCGGAATGATCAAGATGA 1800  
 QY 5353 AAAAATATCAATGTAATGCAAAAGATTAATCTATGACCAATACTATAGCAGTGGAGTT 5412  
 Db 1801 AAAAATATCAATGTAATGCAAAAGATTAATCTATGACCAATACTATAGCAGTGGAGTT 1860  
 QY 5413 GAGAGAGCAAAAGAGCGCTGTCGAAGAGCTTCTGCAAGTACTACCTTGAATTAAGCA 5472  
 Db 1861 GAGAGAGCAAAAGAGCGCTGTCGAAGAGCTTCTGCAAGTACTACCTTGAATTAAGCA 1920  
 QY 5473 GTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAAAGAAATPATGA 5532  
 Db 1921 GTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAAAGAAATPATGA 1980  
 QY 5533 AATAAGAAAAAGCAATGTTAATGTTTCTAGTGAATTAAGAGTCAAGTGTCACAAT 5592  
 Db 1981 AATAAGAAAAAGCAATGTTAATGTTTCTAGTGAATTAAGAGTCAAGTGTCACAAT 2040  
 QY 5593 GCGACAGTCTTTCGAGAGCAAGTGAACAAGCTGAGAGGAGCTGAGTGAAGTGAAT 5652  
 Db 2041 GCGACAGTCTTTCGAGAGCAAGTGAACAAGCTGAGAGGAGCTGAGTGAAGTGAAT 2100  
 QY 5653 AAAATTAACAAAAATCTTCTGCAATTAATAAAAAATAGTAC 5693  
 Db 2101 AAAATTAACAAAAATCTTCTGCAATTAATAAAAAATAGTAC 2141

RESULT 6  
 US-09-841-786-12  
 ; Sequence 12, Application US/09841786  
 ; Patent No. 6669940  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAGARAJA, T. G.  
 ; APPLICANT: STEWART, GEORGE C.  
 ; APPLICANT: NARAYANAN, SANJEEV K.

APPLICANT: CHENGAPRA, M. M.  
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
FILE REFERENCE: 30296  
CURRENT APPLICATION NUMBER: US/09/841,786  
PRIOR FILING DATE: 2001-04-24  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO: 12  
LENGTH: 1887  
TYPE: DNA  
ORGANISM: Fusobacterium necrophorum  
US-09-841-786-12

Query Match 19.4%; Score 1887; DB 3; Length 1887;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1887; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5624 CTGCGATGAGAGCTGGAGTAGACGATTAAATAAAATTACACAAAATCTTCTGCACATATPA 5683  
DB 1 CTGCGATGAGAGCTGGAGTAGACGATTAAATAAAATTACACAAAATCTTCTGCACATATPA 60  
QY 5684 AAAATAGTACTCAAAATGTACGAATGCTTTGGTAAAAAGCAAAATCTCATCTATATA 5743  
DB 61 AAAATAGTACTCAAAATGTACGAATGCTTTGGTAAAAAGCAAAATCTCATCTATATA 120  
QY 5744 AAACAATTGSAATTGGAGCTGGAGTTGGAGCTGGAGAGCTGGAGTGCAGAGTTCTGTAG 5803  
DB 121 AAACAATTGSAATTGGAGCTGGAGTTGGAGCTGGAGAGCTGGAGTGCAGAGTTCTGTAG 180  
QY 5804 CAGTGAATAGATTGTAAATAATACGATAGCAGAAATTAATCATCAAAATCACTGCGA 5863  
DB 181 CAGTGAATAGATTGTAAATAATACGATAGCAGAAATTAATCATCAAAATCACTGCGA 240  
QY 5864 AGGGAATGTCGAGATTATTAACAGATCTGATGCGGTAAATGCTAATTAATGAGAGAACG 5923  
DB 241 AGGGAATGTCGAGATTATTAACAGATCTGATGCGGTAAATGCTAATTAATGAGAGAACG 300  
QY 5924 TGTCTGGAGTGGCCGCTGAGCAATAGAGCCTCAACCAAGTGAATGAATTAACAGAT 5983  
DB 301 TGTCTGGAGTGGCCGCTGAGCAATAGAGCCTCAACCAAGTGAATGAATTAACAGAT 360  
QY 5984 CTACAAAAGCATATGTAAAAAGATTCTACAGTATTGCTAAAGAGAGACAGATGATTATA 6043  
DB 361 CTACAAAAGCATATGTAAAAAGATTCTACAGTATTGCTAAAGAGAGACAGATGATTATA 420  
QY 6044 TTACTACTCAAGGCGCAAGTAGATTAAGTGTAGATTAAGTATTCAAAAATCTTAAATATA 6103  
DB 421 TTACTACTCAAGGCGCAAGTAGATTAAGTGTAGATTAAGTATTCAAAAATCTTAAATATA 480  
QY 6104 AGGAAGCTTATCACAAAAAAGAAAAATAAGTAATAAAAAGAGATTGTTACCAATAGTT 6163  
DB 481 AGGAAGCTTATCACAAAAAAGAAAAATAAGTAATAAAAAGAGATTGTTACCAATAGTT 540  
QY 6164 CAGCTACTCATCTTTAAATCTTTTATTTGGCAAAATGCGCTGGTTCAAGACAGCCGAG 6223  
DB 541 CAGCTACTCATCTTTAAATCTTTTATTTGGCAAAATGCGCTGGTTCAAGACAGCCGAG 600  
QY 6224 TGGCAGGAATCTGTTAATATCAACAGGTTTATGAGAGAACAGAGCTCTTTGTAAGAAAT 6283  
DB 601 TGGCAGGAATCTGTTAATATCAACAGGTTTATGAGAGAACAGAGCTCTTTGTAAGAAAT 660  
QY 6284 CTATATTAATGCAAAACATTAATCTGTAAATCAGAGATTACAGAAATTCATGCGAG 6343  
DB 661 CTATATTAATGCAAAACATTAATCTGTAAATCAGAGATTACAGAAATTCATGCGAG 720  
QY 6344 TAGTAGTTCTGTTGGTGTGTGTAAGATGTAGAGTAGAGCTTCTGTATACCAATA 6403  
DB 721 TAGTAGTTCTGTTGGTGTGTGTAAGATGTAGAGTAGAGCTTCTGTATACCAATA 780  
QY 6404 TTATTAAGAAATACCAAGACAGAGTTGAAAAATCTACATGCTGATGAAGTTTCG 6463

DB 781 TTATTAAGAAATACCAAGACAGAGTTGAAAAATCTACATGCTGATGAAGTTTCG 840  
QY 6464 GAGAGAGCTGAAATTAACAGCAGATTCTAAGCAAGAAATTCCTCTTTTGAAGTCGAG 6523  
DB 841 GAGAGAGCTGAAATTAACAGCAGATTCTAAGCAAGAAATTCCTCTTTTGAAGTCGAG 900  
QY 6524 TGGCAGAGCCGCGGTGGAGCCGAGAGTGGCAGAGACCGTTCCGTAAATCAATTTGAG 6583  
DB 901 TGGCAGAGCCGCGGTGGAGCCGAGAGTGGCAGAGACCGTTCCGTAAATCAATTTGAG 960  
QY 6584 GAAAGCAGAGTAGATGTGAGAGAACCAAGATTTTGGTAAAAAAGCTGAGATTACAG 6643  
DB 961 GAAAGCAGAGTAGATGTGAGAGAACCAAGATTTTGGTAAAAAAGCTGAGATTACAG 1020  
QY 6644 CAATAAGTTATGTTCTGTTGCAATTTGGAATGCCGAGTGGAGTGGCTCCAAAGAG 6703  
DB 1021 CAATAAGTTATGTTCTGTTGCAATTTGGAATGCCGAGTGGAGTGGCTCCAAAGAG 1080  
QY 6704 CTGGAATTGGAGCAGAGCTGGCAGATTACCAAGATGAATCAAAACAGAGCAAGAGTGA 6763  
DB 1081 CTGGAATTGGAGCAGAGCTGGCAGATTACCAAGATGAATCAAAACAGAGCAAGAGTGA 1140  
QY 6764 AAAATTTAAATTAATGACTCGAAAACAGATTAGATGTAATAGCAGAAAAATGAGATPAAT 6823  
DB 1141 AAAATTTAAATTAATGACTCGAAAACAGATTAGATGTAATAGCAGAAAAATGAGATPAAT 1200  
QY 6824 CAGTACTGGAATGCGTTCAAGCCGAGCTGGAATCTTTCAGCCGAGATTTGAGAGTGG 6883  
DB 1201 CAGTACTGGAATGCGTTCAAGCCGAGCTGGAATCTTTCAGCCGAGATTTGAGAGTGG 1260  
QY 6884 TTTCTGTCAATTAATTTGCCAATAAGTAAGAAACAGATGATGATGATCTTAACT 6943  
DB 1261 TTTCTGTCAATTAATTTGCCAATAAGTAAGAAACAGATGATGATGATCTTAACT 1320  
QY 6944 CTTTACTGATGTAATGTAAAAAGCTCTTAATTAATTTGCAATTCCTTGAACGCGGTG 7003  
DB 1321 CTTTACTGATGTAATGTAAAAAGCTCTTAATTAATTTGCAATTCCTTGAACGCGGTG 1380  
QY 7004 GAGGAGCCGAGGCTTTCAGCAGATTAACCGAGTGGTTCTGTAAACATTAATATGTT 7063  
DB 1381 GAGGAGCCGAGGCTTTCAGCAGATTAACCGAGTGGTTCTGTAAACATTAATATGTT 1440  
QY 7064 CTGTATAGCTTGAAGTTCAATACTCTGATTTAATCTCCGTAGAGAAAAAGTAAAG 7123  
DB 1441 CTGTATAGCTTGAAGTTCAATACTCTGATTTAATCTCCGTAGAGAAAAAGTAAAG 1500  
QY 7124 TAAAGCAAAAGAGAAAAAATATTAAGCAAAACAGCAAAATCAGAGATCGAGAG 7183  
DB 1501 TAAAGCAAAAGAGAAAAAATATTAAGCAAAACAGCAAAATCAGAGATCGAGAG 1560  
QY 7184 CAGCAATCGAGCCAAATGCTTGTAAATTAATTTTGGAAACAGCTGTAGAAATGAAAA 7243  
DB 1561 CAGCAATCGAGCCAAATGCTTGTAAATTAATTTTGGAAACAGCTGTAGAAATGAAAA 1620  
QY 7244 ATTCTGAAGAAAAAGAACAGAGTTTAAAACTTTGAAGCAAGTTAACAAAGAACAG 7303  
DB 1621 ATTCTGAAGAAAAAGAACAGAGTTTAAAACTTTGAAGCAAGTTAACAAAGAACAG 1680  
QY 7304 ATTAATAAGTAATGATGCTAGAAAAAATCTTCAATCAGCAGTATTTCTACAGAG 7363  
DB 1681 ATTAATAAGTAATGATGCTAGAAAAAATCTTCAATCAGCAGTATTTCTACAGAG 1740  
QY 7364 ATACTTCTGTAAAGCGAGTAGAGAGATCTCAGGAGAGAGAAATTAAGCCATTGTGA 7423  
DB 1741 ATACTTCTGTAAAGCGAGTAGAGAGATCTCAGGAGAGAGAAATTAAGCCATTGTGA 1800  
QY 7424 AAGCTTCTGATATTAATGGAATAATGTAATTAATCAACAGAGCAAGAAATTAATGA 7483  
DB 1801 AAGCTTCTGATATTAATGGAATAATGTAATTAATCAACAGAGCAAGAAATTAATGA 1860  
QY 7484 CTTTACTGAGTGGTTGGAGCTGCGAG 7510

Db 1861 CTTCTAGTGTGGTGGAACTGCAG 1887

RESULT 7  
US-09-841-786-9  
Sequence 9, Application US/09841786  
Patent No. 6669940  
GENERAL INFORMATION:  
APPLICANT: NAGARAJA, T. G.  
APPLICANT: STEWART, GEORGE C.  
APPLICANT: NARAYANAN, SANGEET K.  
APPLICANT: CHENGARPA, M. M.  
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
FILE REFERENCE: 30296  
CURRENT APPLICATION NUMBER: US/09/841,786  
PRIOR FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: 09/558,257  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 1130  
TYPE: DNA  
ORGANISM: Fusobacterium necrophorum  
US-09-841-786-9

Query Match 11.6%; Score 1130; DB 3; Length 1130;  
Best Local Similarity 100.0%; Pred. No. 9,78-221;  
Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCGGCATCAAAATAACGTCAGAGACAAAGAGAGATATCAGATTCTAAAAA 60  
Db 1 ATGAGCGGCATCAAAATAACGTCAGAGACAAAGAGAGATATCAGATTCTAAAAA 60

Qy 61 GTTTTAATGATTTGGGATGTTGATTAACACTATGACGGGTGAAGGCTTAATGATCAATC 120  
Db 61 GTTTTAATGATTTGGGATGTTGATTAACACTATGACGGGTGAAGGCTTAATGATCAATC 120

Qy 121 ACCCGGATGAGAAATTTTGGAAACAATAAGAAAAAGATTAAGTTTATGACTTACT 180  
Db 121 ACCCGGATGAGAAATTTTGGAAACAATAAGAAAAAGATTAAGTTTATGACTTACT 180

Qy 181 ACAACCAAGATTCAGAGGAGAACGCTTTTACAGTTTAAATAGATTGCTTTAACAGAA 240  
Db 181 ACAACCAAGATTCAGAGGAGAACGCTTTTACAGTTTAAATAGATTGCTTTAACAGAA 240

Qy 241 AATAATATAGCAATCTATATTTTGGGAAAAAGAAATAGTACGGGGGTAAATATCTTTT 300  
Db 241 AATAATATAGCAATCTATATTTTGGGAAAAAGAAATAGTACGGGGGTAAATATCTTTT 300

Qy 301 AACTTGTCAATGAAAAATGAGTAGATGGGATTATCAACGGAATTGCAAAAAATAA 360  
Db 301 AACTTGTCAATGAAAAATGAGTAGATGGGATTATCAACGGAATTGCAAAAAATAA 360

Qy 361 ATTGAGAGAAATTTATTTCTTAAGCTCGAAGGAGATGCGAGTGAAGAAAAATGAGTT 420  
Db 361 ATTGAGAGAAATTTATTTCTTAAGCTCGAAGGAGATGCGAGTGAAGAAAAATGAGTT 420

Qy 421 ATCAATGCTGCTGCTTTTCAATCTATATTTCCAAAAACAAGATGATTTTAAAGAGCTTTG 480  
Db 421 ATCAATGCTGCTGCTTTTCAATCTATATTTCCAAAAACAAGATGATTTTAAAGAGCTTTG 480

Qy 481 GAAAGAGCCAAACATGCTAAAGTTTAAATGGAATCATTCAGTGAATGGAAGTAA 540  
Db 481 GAAAGAGCCAAACATGCTAAAGTTTAAATGGAATCATTCAGTGAATGGAAGTAA 540

Qy 541 ATTCCATTGAATCCGAATGAGAGCATTAAGGTAAGGAAAAATCAATGCTTTGAAGGC 600  
Db 541 ATTCCATTGAATCCGAATGAGAGCATTAAGGTAAGGAAAAATCAATGCTTTGAAGGC 600

Qy 601 ATCGGTTATATATCGGCGGATATTAAGATTGAAGTACTGCAATACATAAGACAGAAAT 660  
Db 601 ATCGGTTATATATCGGCGGATATTAAGATTGAAGTACTGCAATACATAAGACAGAAAT 660

Db 601 ATCGGTTATATATCGGCGGATATTAAGATTGAAGTACTGCAATACATAAGACAGAAAT 660  
Qy 661 ACAGATTTTAAATTTAGTCAATATTAAGTGAATGAATTAATTCGTCGACCGGAGAT 720  
Db 661 ACAGATTTTAAATTTAGTCAATATTAAGTGAATGAATTAATTCGTCGACCGGAGAT 720

Qy 721 TTTAAAGCTACCAAGACAAAATCTGAGATATTAATCTTTGAGCTCACATGATTTCTCT 780  
Db 721 TTTAAAGCTACCAAGACAAAATCTGAGATATTAATCTTTGAGCTCACATGATTTCTCT 780

Qy 781 CAAAAGCTATGGGAAAAAATTCACCTGTTGAAAAGAGAAATGAGAAATATGTAAGGA 840  
Db 781 CAAAAGCTATGGGAAAAAATTCACCTGTTGAAAAGAGAAATGAGAAATATGTAAGGA 840

Qy 841 AATCCAAAGCAAAATATGAAATCTGATGCTGATTTGGAAGAGATGGAATATATAAAAT 900  
Db 841 AATCCAAAGCAAAATATGAAATCTGATGCTGATTTGGAAGAGATGGAATATATAAAAT 900

Qy 901 AGTGGAAAGCTACAAATGGAGATTTATAAAGAAAGGAGAAAAAGAACTTATTAAC 960  
Db 901 AGTGGAAAGCTACAAATGGAGATTTATAAAGAAAGGAGAAAAAGAACTTATTAAC 960

Qy 961 ACTCCTTAAGTTATATCAGATGTGGAAGCTTCGTAAGATTAATAAGAAAAAGTCA 1020  
Db 961 ACTCCTTAAGTTATATCAGATGTGGAAGCTTCGTAAGATTAATAAGAAAAAGTCA 1020

Qy 1021 GAAAAGAAATGTTGACATTAAGTGAAGCAAGAAATTTCTATGATGCAACTTTAGTACT 1080  
Db 1021 GAAAAGAAATGTTGACATTAAGTGAAGCAAGAAATTTCTATGATGCAACTTTAGTACT 1080

Qy 1081 AAGCTTGCAAGCACTCTTTAGCTTTGTTACAGTTCTATTTCTCTAT 1130  
Db 1081 AAGCTTGCAAGCACTCTTTAGCTTTGTTACAGTTCTATTTCTCTAT 1130

RESULT 8  
US-08-973-462-2  
Sequence 2, Application US/08973462B  
Patent No. 6191270  
GENERAL INFORMATION:  
APPLICANT: DRULHE, PIERRE  
APPLICANT: DAUBERSIES, PIERRE  
TITLE OF INVENTION: MALADIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
FILE REFERENCE: 0660-0125-0 PCT  
CURRENT APPLICATION NUMBER: US/08/973,462B  
EARLIER FILING DATE: 1998-02-06  
EARLIER APPLICATION NUMBER: PCT/FR96/00894  
EARLIER FILING DATE: 1996-06-12  
EARLIER APPLICATION NUMBER: FR 95/07007  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 5361  
TYPE: DNA  
ORGANISM: P. falciparum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(5361)  
US-08-973-462-2

Query Match 1.2%; Score 116.2; DB 3; Length 5361;  
Best Local Similarity 41.7%; Pred. No. 8,46-14;  
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

Qy 5126 GTTCGATGAACCGAAGCTTTAGTATGATTCGAGTTGAAGAGTAAATTTCTTCA 5185  
Db 773 GTTGGAAGAAAGATATGCTTCAAGTGTGATGAAGATATGATTCAGATATGAAGAA 832

Qy 5186 ATGATGATGCAAAAGATTAATAATTAATCA---TTGCCGGAATATGCAATGGAG 5242  
Db 833 ATGATGCTCAACATGTTGAAGAAATCGTAGCTCCAAAGTGTGTAAGAAAGTGTGCTCAA 892



Db 3044 AAGAAATCAAGAAATTTATTACAGGTATG-----TTTCGAAGTATAGAAACCACTA 3097  
Qy 7451 TAGATATTACACAGAGACAAGATTAATATCACTTCTACTGCTGCTTGGAACTGCAG 7510  
Db 3098 TAGATATCCATCAGAGAGAAAGGTGATTTGATGAAATATGTTAGTTGATTTTGG 3157  
Qy 7511 GTCTTGCTCCGATCAGAACAGTGGCAGTTACAATATTTAAAGAAATTCGAGTTA 7570  
Db 3158 ATATATATAGAAATATGAAAGAGGTTTATTAATTAATAGAAATATTTCAAGTACTG 3217  
Qy 7571 CTGTTGAAATCTTTTGTGAAAGACGTGAAAGAAAGTAAATCTTATGATCGATAT----- 7625  
Db 3218 AAGGTCTCAAGAACTGTACTGAACATGTAGAACAAATATATATGTGATGTGTATG 3277  
Qy 7626 ----TACAGAAATGTGCTTTACAGCATATCAAGTCTGTAGAGCATTTGGATATG 7681  
Db 3278 TTCCGTCTATGAAATCAATTTTATAGAAATTTAAATGAGCAGAGGTTGAAAGAA 3337  
Qy 7682 GAGCTGCCTATGCAAAATTTAAATTTGAAATGCAATATCAGTATTTAAATTTCTA 7741  
Db 3338 TGTTTATTTATTTGAAAGATGTATTTAAAGTGAAGTATGTATTTACTGTAGAAAGAA 3397  
Qy 7742 AGCTATTAGAGAAAAATATGATGTTATTTGTAATTAATTCGAATTTGAGCGGAG 7801  
Db 3398 TTAAGGATGAACCGGTTCAAAAAAGGTAGAAAAAGAACTGTATTTATTTAGAAAGAA 3457  
Qy 7802 CAAAAGATTTAACCTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAATG 7861  
Db 3458 TGGAGAAAAATTTGTAATGTATTTAGAGAGAAAAAGAAATTTAAACAGCAAGATGA 3517  
Qy 7862 AAATGAATTCAGAGGTGAAATTTGAAAGATATTTTCAATGAGAAAAATAGATTA 7921  
Db 3518 TAGATGCAAGTAAATTCATAGAAATATCTTCAGATTTCTAAAGAAAGAACTGAATCTA 3577  
Qy 7922 GCCCTTTAAAGAAATTTGAAAGAAATCAATGTCAAAGTGAAGAAAGAAACAGAGTA 7981  
Db 3578 TTTAAAGATTAAGAAAAAGTGTCTTCACTAGTGTGAAAGAACTTCAAGCAATGATATG 3637  
Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGCAGTATGACGAGGCGAG 8027  
Db 3638 ATGAAAGTGTGAGAAAGTTTGAATTTGAAATTAATGGAAGAGATTAATGAAGATG 3697  
Qy 8028 AATTAATTCGAGCAAGAAAGATGCGGAAGCTTTATTGAAAGTTAGTACAAAATCCGG 8087  
Db 3698 CTGTTGAATTAATGACATTTACTAGCAAACTTATGAAAGAACTCAAGAGTTAAATGA 3757  
Qy 8088 AAGAGTATTTTTCATGCAGATATGTAATTTGAGACACATTAATGAAGTATAC 8147  
Db 3758 TAGAAGCAGATTTAATTAAGATATGAAATTAATTAAGAAATTTAGAAAAAGC--ATTATC 3815  
Qy 8148 AGCAGTTTCTAAAGCAGTACAGGTGCTGTATTTGGAGAGATTTGAGTCAACCAAGCAGA 8207  
Db 3816 AGAAGATTTCTAAGAAATTAATGATGCAAGATGATACATTTAGAAAAAGTTATTGAAGA 3875  
Qy 8208 AGCTACTGCTGAGGTAAATCTATGTAGAAAGTTGAGAAAGAAATTTGTCAGAACAA 8267  
Db 3876 GGAACATGATTAATACGACGAGTGTGAGAGTGTGAGATTTAAAGATGTGAGAGAGA 3935  
Qy 8268 TCGATGATGCAATTTCTAAGTGAAGGTTTGAATGAAGTAAAGTACGTCTAAATC 8327  
Db 3936 CAAGATCGAAAAAGTATCTGATTTAAAGATTTTGAAGAAATTAATTAAG----- 3988  
Qy 8328 TTCTGTAGTATCAGAAATGAGAGAAATTTGCCGAGCAGAGATGTAATCTTTACAGC 8387  
Db 3989 --AAATTAAGAAATCAAGAACTTGAAAGTAAATTTTGAAGATTTATTAAGAAATTTAA 4046  
Qy 8388 ACAAGTATATCTGAATCCGATGTTGTTACGAAAGCAAGTTATGAAATTAATGATTA 8447  
Db 4047 AACTATGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAAGATCAATTTGAAAA 4106  
Qy 8448 CACAAAAAAATATATTTTGAAGTCAATGCTCTGCTTTAAATGATACAAAGATGAAGC 8507  
Db 4107 ATTGAGAGAAAGCTGAAGAAATTAAGATCTTGAAGCAGATATATTAATTAAGAAAGTATC 4166

Qy 8508 GAATATAGAAATCTTTAGCGGTAGCCGGTGTGCATGCACAAGAAACAAAGCATTTAC 8567  
Db 4167 TTCAATTGAAGTTGAAGAGAAAAAAATTTGAAGAAAGTACACGAATTTAAAGAAAGGT 4226  
Qy 8568 GAGATCAACAGATTTAACTTTACACCTGTAAATGAGAGAAAGTATCTCACTTCGTGC 8627  
Db 4227 AGAATATATTAATGATGTGTATGCGCATATTAAGAGTTTGAAGAAAGTATTTAGAA 4286  
Qy 8628 AAAAGCTTTGGCTAAATATGAAATTTATGAAATGTAAAGAACTGAGAGAGCTTATG 8687  
Db 4287 AGTATGATGA--TTTAAAGAAAGTATATTTAGCATGTTTAAAGGAGATATGGA--ATTAG 4343  
Qy 8688 CGAGCGGAAACACAGCCGTGTAATTTATCAAGAGTACTACAGAGCATTTGTTGC 8747  
Db 4344 GGATATGATTAAGAAAGTTTGAAGATGTAAACAAACAACTTGAAGAAAGTGTGATC 4403  
Qy 8748 AGAATTTGGGAAATTTGAGATTAATTTGAAGATTCAGAGATTAATGATTTGTAAG 8807  
Db 4404 CTTAAGATGTTTATCTAGTATGATGATGATGATGATGATGATGATGATGATGATG 4463  
Qy 8808 AGTCAACGAGACGGAACCAAGAGCTGTCGGAAGAAAGATGATTTCTGTGAAGAA 8867  
Db 4464 AAAAGCTCAAGACCTTAAGTTGAAAGATTTATTTAAAGAAAGGTTAAAGAAAGACC 4523  
Qy 8868 TACAATTTCAAGGAGAAACAAATCATCATTTGAAGATTAAGCCAGAAATTTGGAACCG 8927  
Db 4524 AAAAGAAAAATTAACAAAAAGAAAGTAAAGTTGATTTAAGATTAAGAAACCAAGAA 4583  
Qy 8928 AAGTGTAAATGTAGATGCTTTGAATGAACCTTATGATGATTTCAAGAAAAAGTGTG 8987  
Db 4584 TGAATATGTAAAGTTGAAATGAAAGATGAAAGATGAAAGATGTAAGAAAGATAT 4643  
Qy 8988 CTATGTGGAATTTGTAATTTGAAATGTTGATGTAATTAATGATTAAGAAAAATGTGA 9047  
Db 4644 AGAAGAAATATAGAGAGATTAAGTTGAAGATTAAGATTAAGATTAAGATTAAGAT 4703  
Qy 9048 AGCAAAATCGAAGACATGCTATTTAGAAATCACTGAAACCAAGATTAACAGCATT 9107  
Db 4704 AGGTGAAGACAAAGATGAAAGTTATTAATTAATGTCAAAAAGAAAGCAATTTGAAA 4763  
Qy 9108 TACAAGCAAAAGTAAATTTCTTGAAGAAAGAGACGCTGACAGCTGCAATATC 9167  
Db 4764 GGTTAAGCGAAAAAGAAATTTAGAAAAAGTTGAAAGAGTGTATGCTTTAA 4823  
Qy 9168 GAATGTACACTTTCATGATGATGATTAATTAATTTGCAAGCAGTATGATCTTC 9227  
Db 4824 AAAACAGTAGACAGATTAATGAATATGTTCAAAAAATGATTAAGAAAGTTGATTAAGA 4883  
Qy 9228 TCAATTAATAACCAAAATTCAAAAATTAATATTACTTTAGCATCAA 9274  
Db 4884 AGTATCTTAAGCTTTAGATCAAAAAATGATTAAGTACTTAATGTTTAA 4930

RESULT 9  
US-08-973-462-1  
; Sequence 1, Application US/08973462B  
; Patent No. 6191370  
; GENERAL INFORMATION:  
; APPLICANT: DRULHE, PIERRE  
; APPLICANT: DAUBERSIES, PIERRE  
; TITLE OF INVENTION: MALALIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
; FILE REFERENCE: 0660-0125-0 PCT  
; CURRENT APPLICATION NUMBER: US/08/973.462B  
; CURRENT FILING DATE: 1998-02-06  
; EARLIER APPLICATION NUMBER: PCT/FR96/00894  
; EARLIER FILING DATE: 1996-06-12  
; EARLIER APPLICATION NUMBER: FR 95/07007  
; EARLIER FILING DATE: 1995-06-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 6152

TYPE: DNA  
; ORGANISM: P. falciparum  
US-08-973-462-1

Query Match 1.2%; Score 116.2; DB 3; Length 6152;  
Best Local Similarity 41.7%; Pred. No. 8.8e-14;  
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

QY 5126 GTTCGATGAAACGAGAGCTTTAGTTAGATTGAGATTGAGAGAGTAAGTCTTTCA 5185  
DB 1017 GTGTTGAAGAAAGTAGCTTCAAGTGTGAGAAAGTATGATTCAGATATTGAAGAA 1076  
QY 5186 ATGTGATGCAAAAGATCAAAAACAATAATCA---TTGCCGAAATGCAATGAG 5242  
DB 1077 ATGTAGCTCCAACTGTGAAAGAAATGTAGCTCCAAAGTGTGAGAAATGTGCTCAA 1136  
QY 5243 GAAAACGCGCTGAGATTGAGCAACAGTGTCTATACAAATATTGAAAAAATCAATGTA 5302  
DB 1137 GTGTTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTAGCTGAAATGTTGAAGAA 1196  
QY 5303 TAGCTATTGTAAAAACAGTAAATTTACAACGCGCAATGATCAAGTAGAAAAATATCA 5362  
DB 1197 GTGTAGCTGAATATGTTGAAGAAAGTAGCTGAATATGTTGAAGAAAGTAGCTGAAA 1256  
QY 5363 ATGTGATGCAAAAGATTACTATGACCAATATATAGCAGTCGAGTTGAGAGCA 5422  
DB 1257 ATGTGAAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAA 1316  
QY 5423 AAGGAGCTCTGTGCAAGAGCTTCTGCAAGTACTACTTGAATAGACAGTTCTTCTC 5482  
DB 1317 TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA 1376  
QY 5483 ATGTGATCAAACTGATATTGACAAAGATTGAGAGAGAAATTAAGAAATAGGAA 5542  
DB 1377 ATGTGAAAGAAAGTAGCTGAAATGTTGAAGAAAGTAGCTGAAATGTTGAAGAA 1436  
QY 5543 AAGCAAAATGTAATGTTTCTAGCTGAAATACAGTCAAGTGTCAAAATGCGACAGTGC 5602  
DB 1437 GTGTAGCTAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAA 1496  
QY 5603 TTTCCGAGCAAGTGCACAAGCTGCAGTAGAGCTGAGTAGCAATTAAATTTACAC 5662  
DB 1497 ATGTGAAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAA 1556  
QY 5663 AAAATATCTTGCACATATPAAAAATAGTACTCAAAATGTACAAATGCTTTGTGTA 5722  
DB 1557 TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA 1616  
QY 5723 GC---AAATCTCATTTATTTAAACAATTGGAATTGAGCTGGAAGTTGAGCTGAG 5779  
DB 1617 ATGTGAAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1676  
QY 5780 GAGCTGAGTGAACAGGTTCTGTAGCAGTGAATAGATTGTAAATTAATTCAGTAGCAAT 5839  
DB 1677 GTGTAGCTGAATGTTGAAGAAAGTGTAGCTGAATATGTTGAAGAAAGTGTAGCTGAAA 1736  
QY 5840 TAAATCATGCAAAATCACTGCGAAGGAAATGTGCGAGTTATTACAGAGCTGATCGG 5899  
DB 1737 ATGTGAAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1796  
QY 5900 TAATTGCTATTATGACAGAACAGTGTGGAAGTGCCTCGACAGAAATGAGAGCTCAA 5959  
DB 1797 TCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTTAGCTCAA 1856  
QY 5960 CCAAGTGAATGAATTTACAGATCTTACAAGCATATGTAAAGATTCTACAGTAGTG 6019  
DB 1857 GTGTTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAAGTGTGAAGAA 1916  
QY 6020 CTAAAGAGAAACAGATGATTATTTACTACAGGCGCAAGTAGATTAAGTGTAGATA 6079  
DB 1917 GTGTAGCTGAATATGTTGAAGAAAGTGTAGCTGAATATGTTGAAGAAAGTGTAGCTGAAA 1976  
QY 6080 AAGTATTCAAAATCTTTAATTTATACGAACATTATCCAAAAAGAAAAATAGTAATA 6139

DB 1977 ATGTGAAAGAAAGTAGCT-----GAAATGTGAAGAAATCGTAGCTCCAACTGTTG 2030  
QY 6140 AAAAAGATTTGTATCCAAATAGTTCAGCTACTCATACTTTAAATCTTATTGCAATG 6199  
DB 2031 AAGAAATGTGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAAAGTGTGTA--GAAAGTG 2087  
QY 6200 CCGCTGTTTCAAGCAACCGGAGTGCAGGAACTGTTAATATCAACAAGTTTATGAG 6259  
DB 2088 TGGCTTCAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAAATG 2147  
QY 6260 AAACAGAACTCTGTGAAAAATCTATATTAATTAATGCAAAACATTATCTGTAATATCAG 6319  
DB 2148 TTGAAGAAAGTGTAGCTGAATATGTTGAAGAAAGTGTAGCTGAATATGTTGAAGAAATCG 2207  
QY 6320 GAGATTACCAATTCATTCGAGTAGTAGTCTGTGTGCTGTGTGTGTGTGAAATGTAAGAG 6379  
DB 2208 TAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAAAGTG 2267  
QY 6380 TAGAGCTTCTTGTGATCCAAATATTTAATAAG---AAATPCCAAAGCAAGATTGGA 6436  
DB 2268 TTGTGAAGAAAGTGTGCTCCAACTGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTG 2327  
QY 6437 AAATCAATATGTGTGATGAAGGTTTCGAGAGAAAGCTGAATTAACAGCAATTCTAAGC 6496  
DB 2328 TAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAATATGTTGAAGAAAGTGTAGCTGAAAAATG 2387  
QY 6497 AAGAAATTTCTCTTTTGAAGTGTGAGTGTGCGACGACCGGGGTAGAGCCGAGTGTGCGAG 6556  
DB 2388 TTGAAGAAAGTGTAGCTGAATATGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATCG 2447  
QY 6557 GAACGCTTTCGTAATCAATTTGCAAGAAAGCGAAGTAGATGTGGAAGAACAAAGA 6616  
DB 2448 TAGCTCCAACTGTTGAAGAAATGTTAGCTCCAAAGTGTGTAGAAAGTGTGCTCCAAAGTG 2507  
QY 6617 TTTTGTAAAAAAAGCTGAGATTACAGCAAAACGTTATAGTCTGTGCAATGTGAATG 6676  
DB 2508 TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAATATGTTGAAGAAAGTG 2567  
QY 6677 CCGCAGTGCAGTGTGCAAAAGAGCTGGAATTTGAGACAGCAGTGCAGATTACCAAG 6736  
DB 2568 TAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTCCAACTG 2627  
QY 6737 ATGATTCAAACACAGAGCAAGTGAAGAAATTTCTAAAT--ATGATCCAAACAACT 6793  
DB 2628 TTGAAGAAATGTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGTG 2687  
QY 6794 TAGATGTAATAGCAAAATGATTAATCAGTACTGGAATCGGTTCCAGCCGAGCTG 6853  
DB 2688 TTGCTGAAGAACTTCCAAACAAATTTATCAGACAACTTTTAAGTAATTTAAGGTGTA 2747  
QY 6854 GAATCTTGCAGCCGAGATATCGAGTGTCTGTCAATATATTTGCAATTAAGTATG 6913  
DB 2748 TCGAAACTGAGGAAATTAAGGACAGATATTTAAAGAGTAGAAGAAAGTGAAGAAATG 2807  
QY 6914 AAACAGATATGATATAGTACTTTACACTCTTCTACTGATGTAATGTAAGTAAAGCTCTTA 6973  
DB 2808 TAGTCCCAACAAATACAGAAAGCTGAAGAAACATACAGCTGAAGTGTATCTCTTTTA 2867  
QY 6974 ATPAAATTTGCAATTTCTTGACAGCCGCTGAGAGACCCGAGTCTTGCAGCAGTTACCG 7033  
DB 2868 GTTACATATTAAGAGGATACAAAGAAATATCATTAATGATTAATAGAGGAAATAT 2927  
QY 7034 GAGTGTCTTCTGTTAACATATTAATATGTTCTGTATAGCTGAGTTCACATTAATCTCG 7093  
DB 2928 TAGAAGAACTCCACGAAATGTAATTAAGTCCGCTTTAAGAAATATCCAAAGTGAAGAG 2987  
QY 7094 ATTGACTTCCGTAAGAG---AAAGTAATGTAAACGCAAAAGGAAAAATATATTA 7150  
DB 2988 AAAAGAAAGATATGATGTTATGTAAGTGAAGAGTGAAGAAAGTGTGCTACCACTTTAA 3047  
QY 7151 AGCAAAACAGCAATATGACAGAAATCGAGAGCAGCAATCGAGCCAAATGCTTGTGTA 7210

Db 3048 TAGAACTGTGGACAGGACAGAGAAAAGCGCAATATCAATTACGAAATATTGGAAA 3107  
Qy ATAAATTTGGAAACAGCTGTAGAGATPAGAAAAATTTGAGGAGAAAAGAAACAGAGTTT 7270  
Db 3108 ATTTGAAAGAAATGAGTAGAAGTATGAAAAATGTTGACAGAAATTTTGAAGAAATTA 3167  
Qy 7271 TAAAACTTTAGACGAAAGTTAACAAAGAACAGATAAAAAGTAAATGATGCTACGAAA 7330  
Db 3168 ACGAACTGTATTTAACTGATTAAGATPAAAGTAGAGAAACAGTAGAAATTTACCGAG 3227  
Qy 7331 AAATCTTACATCAGCAGATTTCTTACAGAAAGATCTCTGTAAAGCGATAGAGAG 7390  
Db 3228 AAAAGTTGAAAAACAATGAAATGATPAAAGCATTTTTAGTAAATATTGTGTA 3287  
Qy 7391 ATACTCAGGAGAGAGAAATTAAGCCATGTGTGAAGCTCTGATTTATTTGAAAAAATG 7450  
Db 3288 AAGGAATACAGAAAAATTTATTAACAGTATG-----TTTGAAGTATAGAAACAGTA 3341  
Qy 7451 TAGATATTAACAACAGAGACAAAGAAATATATCACTTCACTGTGTGTGGAACTGCGAG 7510  
Db 3342 TAGTAAATCCAAATCAGAGAAAAGTTGATTTGAATGAAAAATGTGTTAGTTGATTTAG 3401  
Qy 7511 GTCTGTCTCCGCATCAGAGACAGTGCAGTTTCAAAATTTAAAGAAATTCGAGATTA 7570  
Db 3402 ATAAATATAGAAAAATAGAAAAGAGTTTATTAATTAATTAAGAAAAATATTTCAAGTACTG 3461  
Qy 7571 CTGTGAAAAATCTTTTGTGAAGACAGCTGAAAAAGTAAATGTATGATCGAGTAT----- 7625  
Db 3462 AAGGTGTTCAAGAACTGTACTGAACATGTAGAACAAATATATATGTGGATGTTGATG 3521  
Qy 7625 ----TACAGAAATTTGCTTTTACAGCATCAAGCTCTGTAGAGACATTTGGAAATAG 7681  
Db 3522 TTCCGTATGAAAGATCAATTTTAAAGAAATTAATAGACAGAGAGGTTGAAAGAAA 3581  
Qy 7682 GAGCTGCTATGACAGATTTAAATTTCTAATGGAAGTCAATATCGATTTAAAAATTTCTA 7741  
Db 3582 TGTTTTTAATTTGAAAGATGTATTTAAAGTAAAGTGAATGTAAATTTCTGTAGAGAAA 3641  
Qy 7742 AGCTATGAGAAAAAATATGTATGTATTTGTAAGAAATTAATCGGAATTTGAGCGGAG 7801  
Db 3642 TTAAGATGAAACCGGTTCAAAAAAGGTAAGAAAAAGAACTGTAGTATTTAGAGAAA 3701  
Qy 7802 CAAGAAAGATTAACGTAAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAGAAATG 7861  
Db 3702 TGGAGAAAAATATGTATGATGTATTAAGAGAGAAAAAGAAAGATTTAACAGCAAGATGA 3761  
Qy 7862 AAATGATTTCAAGGTTGAAATTTGAGAGAGATTTTCAATGAAGAAAAATAGATTACTA 7921  
Db 3762 TAGATGCAGTAGAAATTCATAGAAATATCTTCAGATTTCTAAAGAAAGAACTGATCTA 3821  
Qy 7922 GCCCTTTTAAAGAAATTTGAGAGAGAAATCAATGTCAAAAGTGGAAAAAGAAACAGATGA 7981  
Db 3822 TTAAGATTAAGAAAAAGATGTTTCACTAGTTGTGAGAGAGTTCAAGCAATGATATAG 3881  
Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGACAGTACAGGCGGACAG 8027  
Db 3882 ATGAAAGTGTGAGAAAGTTTAAAGATTAAGAAAAATATGGAAGAGGTTTAATGAAGATG 3941  
Qy 8028 AATTAATTTCCGAGCAAGAAAGATGCGGAAAGCTCTTAATTTGAAAGTTAGTACAAAATCCGG 8087  
Db 3942 CTGTGAAATTAATGATCTTACTAGCAAACTTATTTGAGAAAACTCAAGAGTTTAAATGAG 4001  
Qy 8088 AAGAGTATTTTTCATGACGATATGTGAATTTGAGAGCAACATTAATGAAGTATAC 8147  
Db 4002 TAGAAGCAATTTAATTAAGATATGAGAAAAATTAAGAAATTTAGAAAAAGC--ATTATTC 4059  
Qy 8148 AGCAGTTTCAAGACAGTAAAGAGTCTGTATTTGGAGAGATTTGAGTCCAGCAAGCGAGA 8207  
Db 4060 AGAAGATTTCTAAAGAAATATAGATGCAAAAGATGATACATTAGAAAAAGTTATTGAGAA 4119  
Qy 8208 AGCTACTGCTGACAGGTAATCTATGTAGAGATTTGAGAGAGAAATTTGTTCAAGAACAA 8267  
Db 4120 GGAACATGATATTAACGACGACGTGTGATGAAGTTGTAGAAATTTAAAGATGTGCAAGAGAA 4179

Qy 8268 TCGATGGAATGCAATTTCTAAAGTAGAAGTTTGGATGATGAATTAAGTAATCTGCTAAATC 8327  
Db 4180 CAAGTCGAAAAAATATCTGATTTTAAAGATCTTGAAGAGATATATTTAAAG----- 4232  
Qy 8328 TTCTGTAGTATCAGAGAAATGAGAGAGAAATTCGCGAGCAGAGTGAATCTTACAGC 8387  
Db 4233 --AAGTAAAGAAATCAAAAGAACTTGAAGAGAAATTTTAAAGATTTTAAAGAAATTTAA 4290  
Qy 8388 ACAAGTAAATCTGAATCCGATGTTCTGTTACGAAGCAAGATTAATGAAATTAATGATTA 8447  
Db 4291 AACTATGAAACAGATATTTTGAAGAGAAAAAGAAATTAAGAAATGATCAATTTGAAAA 4350  
Qy 8448 CACAAAAAATATATTTTCAAGAGTCAATGCTCTTGTCTTAAATGATACAAAGAAATGAAGC 8507  
Db 4351 ATTGGAAGAAAGAACGTAGAGAAATTAAGATCTTGAAGCAATATATTTAAAGAAAGTATC 4410  
Qy 8508 GAATATAGAAATCTTTAGCGTAGCCGCTGTGATGACACAGAAACAAACAAAGCATTTAC 8567  
Db 4411 TTCAATTGAAAGTGAAGAAAGAAAAAATTTGAAGAAATGACAGAAATTTAAAGAAAGGT 4470  
Qy 8568 GAGATCAAAACAAGTTAATCTTACAACTGTAAATGAGAGAAACGATCTCACTTCGTC 8627  
Db 4471 AGAACATATATTAAGTGTGATGCGCATATPAAAGTTTGAAGAGATGATTTAGAGA 4530  
Qy 8628 AAAAGCTTGGCTAAAAATGAAAAATTAATGAAATGTAAAAAGAACTGGAGAGAGCTTACT 8687  
Db 4531 AGTATATGA--TTTAAAGAAAGTATATTAAGCATGTTTAAAGAGATATGGA--ATTAG 4587  
Qy 8688 CGAGCGGAAACAGCAGCCGTTGAAAAATTAATCAAAAGATCTACAGAGCATTTGTTGC 8747  
Db 4588 GGATATGATATGAGAAAGTTTGAAGAGTATTAACAACAAACCTTGAAGAAAGTTGAATC 4647  
Qy 8748 AGGAAATTTGGAAATTTGAGATTAATTAAGAAAGATTTGCAAGATTAATGATTTGTAG 8807  
Db 4648 CTTAAAGATTTTATCTAGTGAATTTAGCATGATGATGAAGAAACAAATGAAACAGAAA 4707  
Qy 8808 AGTCAACGAGACCGAAACCAAGAGAGCTTGTGGAAGAGATGATTTCTGTGAAAAA 8867  
Db 4708 AAAAGCTCAAGACCTTAAGTTGGAAGAAAGTATTTTAAAGAAAGAGTTTAAAGAAACC 4767  
Qy 8868 TACAATTTCAAGGAGAAACAAATCATCCATGGAAGATTAAGCCAGAAATTTGGAAACCG 8927  
Db 4768 AAAAGAAAAAATTAACAAAAAGAAAGTGAAGTTGATTTTAAAGATTAAGAAACCAAAA 4827  
Qy 8928 AAGGTATATGATGATGCTTGAATGAACCTTGAATGATCTACAGGAAAAAGTGTG 8987  
Db 4828 TGAATATGATAGAAAGTTGAAATGAAAGATGAAGATATGAAGAGATGTGAAGAAAGATAT 4887  
Qy 8988 CTATGTTGAAATTTGTGAAAAATGTTGATGTAATATGATTTAAGAAAAATGTAGA 9047  
Db 4888 AAGAAAGATATTAAGAAAGATTAAGTTGGAAGATATGAATGAATATATATGATGAAGATAT 4947  
Qy 9048 AGGTAAGACAAAGATGAATTAATGATTAATGTCCAAAAAGAAACCATTTGAAAAA 5007  
Db 4948 AGGTAAAGCAAAAGAAAAATTTAGAAAAAAGTTGAAAGAGTGTAGTGTCTTAA 5067  
Qy 9108 TACAAGACAAAGTAAATATTTCTTGAAGAAAGAGACGCTCAAGTGAAGTGCATATC 9167  
Db 5008 GGTTAAGCCGAAAAAGAAAAATTTAGAAAAAAGTTGAAAGAGTGTAGTGTCTTAA 5127  
Qy 9168 GAATGTACAAATTTCAATGATGATATTAATTAATTTTGGCAAGACGATATGATCTTC 9227  
Db 5068 AAAACAGTATGACGAGATTAATGAATATGTTCCAAAAATTTATATTAAGAGATTTAAAG 5127  
Qy 9228 TCAATTAATTAACAAAAATTTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTA 9274  
Db 5128 AGTATCTTAAGCTTTAGAAATCAAAAAATGATGTTACTATATGTTTAA 5174

RESULT 10  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463

Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZgpt-F1s  
US-08-232-463-14

Query Match 0.9%; Score 87.6; DB 2; Length 7218;  
Best Local Similarity 2.4%; Pred. No. 6.3e-08;  
Matches 9; Conservative 251; Mismatches 120; Indels 0; Gaps 0;

QY 7098 GACTCCGTACGAAAAATGTAATGTAACGCAAAAGAGAAAAATATTAGCAAC 7157  
DB 1445 GATTGTTGTAACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1386  
QY 7158 AGCAGCAATGACGAGAGAGCAATCGAGCAATCTGCTGTAATAATT 7217  
DB 1385 RRR 1326  
QY 7218 TGAACAGCTGTGAAGATGAAAAATCTGAAGAAAAAGAAAGCTTTAAAG 7277  
DB 1325 RRR 1266  
QY 7278 TTAGAGCAAGTAACAAGAACAGATAAAAAGTAATGATGTAAGAAAAATCTT 7337  
DB 1265 RRR 1206  
QY 7338 ACAATCAGCAGTATTCTACAGAGATACTTGTAAAAAGCGATAGAGAGATACT 7397  
DB 1205 RRR 1146  
QY 7398 GGAAGAGAAATTAAGCCATTGTGAAGACTTGTGATATTATGAAAAAATGATAT 7457  
DB 1145 RRR 1086

QY 7458 TACAACAGAGACAGAGATA 7477  
DB 1085 RRRRRRRRRRRRRRRRRRR 1066

RESULT 11  
US-08-956-171E-63  
Sequence 63, Application US/08956171E  
Patent No. 6593114  
GENERAL INFORMATION:  
APPLICANT: Charles Kunesh  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5256  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8155 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 63:  
US-08-956-171E-63

Query Match 0.8%; Score 79.2; DB 3; Length 8155;  
Best Local Similarity 41.3%; Pred. No. 3.4e-06;  
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

QY 3782 TTCAAGTAAACAAGTCTTTGGTTAAATAAGATGATTTGGAATGCAATTAATT 3841  
DB 5651 TTAATTAACCACTGACGTAATGAAATCAGCAATTTTAATTAACAATTGCAAG 5710  
QY 3842 AAGTAAGCTTGAAGTGAAGTCAAGTACAGAGAGAGAGAGAGTGTGAAGCAGT 3901  
DB 5711 AATTCAGCTACGCAAGTGAAGTGAAGAAAAACAAGCAGTATGCTGAAGCAA 5770  
QY 3902 AAGAAAGTGAGACAGAAAGAAAGTTATCTATTGGAACTTCTGCTTATCACTT 3961  
DB 5771 ATACTGAATAATGTAAGCAATCAAGCAATTCAGCAGCACTAAGCAAGTGG 5830  
QY 3962 TGAACATGAATTTCTGCAAAATCAGAAATTAATCACTAGCAGAGAAATGAAAGCC 4021

Db 5831 ATGAAGCTAAAGCAATGCAAGACGCGATTAAAGCGTTAAACACCAAAAGTTGGAAGA 5890  
 Oy 4022 AAAAAATGATGTGATGCTCACTGCTTATCAAGCGGACACCCAGATGACAGGAGCTTAA 4081  
 Db 5891 AACCAAGCGCTAAAGATGAATTTGATTCATTTACAAAGCAACGCAACCAATGTTATCAAT 5950  
 Oy 4082 ATTTCAGAGCTGGAAGTCMAATGGAATGTAGGGGCTACTGTGACTGTGGCCAAATTTA 4141  
 Db 5951 ATGATCAGAACGCTCAACAGAGAAAAGAAAGACGCTATTCAACAAATTAGCAACGACG 6010  
 Oy 4142 ACACAAAGTAAATGCTTCTATTAGTGTGGAGATTAATTAAGTTAATGACGCGACG 4201  
 Db 6011 TTACAGACGCGCAAAATATATATTACAGCTGCAACTGATGATATGTGATGTCAGGCGA 6070  
 Oy 4202 CAAAAGCTCTTTTACCAACCACTCAAGTGACTGTGACGTGACGACGGGAGGACAAATTA 4261  
 Db 6071 AAGACGCTGGAAGAAATTCATTTCAAGACGCAACGCAACGACGCGTTAAATCAATG 6130  
 Oy 4262 GTTCTGAGCGGGAATTAAGAAATTAATCAAGGGGCTGTTCTGTCAATAGATTGACAAATG 4321  
 Db 6131 CTAAAAATGATGTGATCAAGCTGACAACTCAAAATCAAGCAATTTGATATACACTG 6190  
 Oy 4322 AGGTGAAGCTAGCGCTTGATTAATCTTCATGGAAGAGCTAATGAAATCAATGTCATTTG 4381  
 Db 6191 GTGTACACACTGAAAGAAAATGACAGAAAAGATTGTTTAAAGCTAAAGAAAAG 6250  
 Oy 4382 CCAAAAGTGTCAAGGAAGTTGATCTGATCTGCAAAAAGAAATACAGGCTTTACTAAATGGA 4441  
 Db 6251 CGTATCAAGATATCTTTAAATGACAAACACTAATGATGTTAGCAAAATTAAGATTCAG 6310  
 Oy 4442 AAGATAAAAATATTAGAAAGATCGTGTATTATACAGACTGGAATGGTTATTTATACGA 4501  
 Db 6311 CAGTGTGATATTCAAGGATATTACTGACGATACAAACATTTAAAGATGTTG-----CGA 6364  
 Oy 4502 AAGAACACTAGAAAAAGCAAAAAGAAAAGAGGCGGCTCATTTGTTAAATGCTGCTTAT 4561  
 Db 6365 AAGATGAATTAGCAACAAAGCAACGAAACAAAAGCGCTTATTCACAAACTGACAGATG 6424  
 Oy 4562 CGGTGCTGGAACGATAAATCGGCTGAGAGAGTACTATTGCACTCAATCTGTTAAA 4621  
 Db 6425 CGACTACTGAAAGAAAAGAAACAGCAATCAACAGTGAACGCAACAAATTAACACAGGTA 6484  
 Oy 4622 ATAAATTTAAAGCAATTTAGTGAAGCAATTAAGAAACCGAGAGAGATAAATTCATG 4681  
 Db 6485 ATCAAAATATTGAATGCAAGTCAATCGATGATGTAACACTGCCAAAAGTATGCA 6544  
 Oy 4682 CGAAACATGTAATGTGAGGCAAAATCATCTACTGTTGTTGTGATGCGGCTTGAC 4741  
 Db 6545 TTCAAGCAATTGACCAATTCAAGCATCAACAGATGTTAAAACGATATGCAAGAC-GGA 6603  
 Oy 4742 TTGCTATCAGCAAAAGATGCTTTTCAAGGAATGGGATCTGGAGCATGGGACCTTATCA 4801  
 Db 6604 TTGCTAATCTGAATCAAAATTAATAAATACTGAATACTTAATTAATGAGCT----- 6657  
 Oy 4802 ATGACACGATTTGCAAGGTGATTAAGGAAGAAATTTCTGCTGATTCCTTAATGTGAACG 4861  
 Db 6658 -----ACTAATGAAGAAAAGTTACGATATTGACAGCTTAAGACAGATATGAAGAG 6712  
 Oy 4862 CAATTAATTCATTTCTTGAGGATGAATGTTGCGGAAACATTCGCGTTCTCTTTACGG 4921  
 Db 6713 GTTAAATTAATTTATGACAGCACTACTACAGGTGATGTAATCTCTGTAAGATACAG 6772  
 Oy 4922 CGGTAGGAGCTGCTTTTGCAATTAATTAATCTTCATTAATTAACATCTGCTTTGATTAAG 4981  
 Db 6773 CAGTCAAAAAGATTCAACAACTTCATGCAAACTCTGTTAAGAAACAGAGGTTAAAAAG 6832  
 Oy 4982 GAACGAAGTAAATCTTTTATGTAAGAAAGATACAAAAGTCAATGTACAAAGCTTTGAATG 5041  
 Db 6833 AATTGATCAAGCTCAGCTGATTAAGAAAACAAATGGAACAAACCAATATGATCAC 6892  
 Oy 5042 ATTCTCATTTTCAAAAGCTTTCTGCTGAGGCGCTGCAAGTATTAAGACAGCTGGAATCG 5101

Db 6893 AACAAAGAAATTAATGATGCAAAACAGAAAGTTGATTAAGTAAATCAACGCAAAACAA 6952  
 Oy 5102 GAGGAATGATCTGTCAATCGTGGTTCTGATGAAGAAAGCAAGCTTATGATGATTCG 5161  
 Db 6953 ATGTGATCAATTCATCAACCAATGAATATGTTGATTAATGCAAGTTAAAGAGAAAAGCTA 7012  
 Oy 5162 AGTTGAAGAGTAAGTTCTTTCAATGTAATGCAAAAAGATCAAAAACAAATAATACAA 5221  
 Db 7013 AAATTAATGAGTTAAAAATTTAGTGTGCAAAAAAGATGCTTTACTAAATTTGAAG 7072  
 Oy 5222 TTGCGGAAATGCAATGAGAGAAAAGGCGTGAAGTTGAGCAACAGTGTCTATCA 5281  
 Db 7073 ATGATATTAATGCTAAAGATTAACAGAGGATTAATCTTAACGATCGACTTCAAGTGAA 7132  
 Oy 5282 ATATTGAAAAATCAATGATTAATGATTAATGTAATTAATTAATTAATTAATTAATTAAT 5341  
 Db 7133 TTGCTGAAGCAAAACAAAATCTTGTAATTTAAACAACTGCGGATCAAAATGTTAATC 7192  
 Oy 5342 ATCAAGATGAAAAAATATCAATGTGACTGCAAAAAGATTAATTAATGACCAATTAATG 5401  
 Db 7193 AAGCTACTTTAAAGATGACATTGAAGTTCAAAATTCATTAATGACTTAATTAATTAACG 7252  
 Oy 5402 CAGTCGAGTTGAGAGCAAAAGAGCCTCTGTGCAAGAGCTTTGCAAGTACTACT 5461  
 Db 7253 ATTAACAAATTCACACAGTTAAAAAGAAATGAGCTACACAGATTTATATGCTTAATG 7312  
 Oy 5462 TGAATTAAGACAGTTCTCTCATGTTGATCAAACTGATTAATGACAAAGATTTAGAGAG 5521  
 Db 7313 ATCAAGAAAAAATTAATTTTCACTGACCTAATGCAACACAGATTAATTAATTAATTAAT 7372  
 Oy 5522 AAAAAATGAAAAATPAAGAAAAAGCAATGTTAAATGTTTCTGCTGAAAAATGAGTCAAG 5581  
 Db 7373 CAATTAAGCAAGTTGACCAAAATG---TTCAACTGCTATTGAAGCAATTAAATATGCTG 7429  
 Oy 5582 TGTGTCAAAATGCAACAGTGTCTTCCGAGCAAGTGAACAGTGCAGTACAGTGTGAG 5641  
 Db 7430 TGATTAATGAGTGAAGTGTGATGATGATTAACAAAGGTAAGGCAAGTATGATGTTATTC 7489  
 Oy 5642 TAGCAGTAATTAATTAATTAACAAATTAATCTGACATATTAATTAATTAATTAATTAAT 5701  
 Db 7489 AAGTGAATGCTAATCTTTAACTTAAGGCAACAGCTAATGAAATTAAGCAAGATA 7549  
 Oy 5702 TAGCAATGCTTTGTTAAAAAG-----CAATCTCATTCATCTA 5740  
 Db 7550 CGAAAGATCTAATGATCAAGTGCACAGTTAACTGCTGAAGAAAACTGAAGCATTAAG 7609  
 Oy 5741 TTTAAACATTTGAAATTTGAGCTGAGTGTGAAGCTGAGAGAGCTGAGTGCAGGTTTG 5800  
 Db 7610 CAATGATTTAAACAAATTAACAGATCAAGCTAAACAGGATTAATCTGATGCAACCAACTG 7669  
 Oy 5801 TAGCAGTGAATTAAGATTTGTAATTAATACGATAGCAGAAATTAATCATGCAAAATCACTG 5860  
 Db 7670 CTGAAGTTGAAAAAGCGAAAGCTCAAGGACTTTGAAGCAATTTGATTAACATTCAGACT 7729  
 Oy 5861 CGAAGGAAATGTCGAGATTTAACAGAGTGTGATGCGTAATTTGCTAATTAATGACAGAA 5920  
 Db 7730 CAACAGAAAAACAAAAAGCTATGAGAAATTAAGAACTGACCTGACCAAGATTGAAGCAG 7789  
 Oy 5921 CAGTGTCTGAGTGGCCCGTGCAGCAATAGAGGCTCAACAGTGTGAATGAATTAACAG 5980  
 Db 7790 GTGTAAATGTCAGCTGATGCTACAACTGAAGAAAAAGAGGCTTAAGATGCTTTAG 7849  
 Oy 5981 GATCTACAAAAGCATATGTAATAAGATTTACAGTATTTGCTAAAGAAAGAACAGATGAT 6040  
 Db 7850 AAGACATTTTATCAAAAAGCACTGAAGATATTTCTGATCAAACTACAAATGCAAAATG 7909  
 Oy 6041 ATATTACTACTCAAGGCAAGTATGAATTAAGTGTGATTAAGTATTCAAAAATCTTAATA 6100  
 Db 7910 CTACTGTCAAAAATAGTGGCTTGAAACAATTAAAGCAACGATTAATTAATCTGAAGTTA 7969  
 Oy 6101 TTAACGAAGCTTATCACAAAAAAGAAATTAAGTATTAATAAAAGATTTGTTACCAATA 6160  
 Db 7970 AGAAAAATGCTTTGGAAGCAATCAGAGAGTGTTAACAAAGCAATAGGAATTAATTAATA 8029

QY 6161 GTTCACTACTACTT 6178  
DB 8030 ATGCAGATGCAGATGCAT 8047

## RESULT 12

US-08-781-986A-63  
Sequence 63, Application US/08781986A  
Patent No. 6737248

## GENERAL INFORMATION:

APPLICANT: Charles Kunach  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255

## CORRESPONDENCE ADDRESS:

ADDRESS: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA

ZIP: 20850

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ. ID NO: 63:

SEQUENCE CHARACTERISTICS:

LENGTH: 8155 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-781-986A-63

Query Match 0.8%; Score 79.2; DB 3; Length 8155;  
Best Local Similarity 41.3%; Pred. No. 3.4e-06;  
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

QY 3782 TTCAAAGTAAACAAGTCTTTGGTTAAATAATGATATTCGAAATTCGCAATTAATTAA 3841  
DB 5651 TTATATTAATAAGCACTGCGATGTAATGAATTCACAGCAATTTAAATTAACAATTGCAAG 5710  
QY 3842 AAGTAAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGCAGGTTTGAAGCAGTTA 3901  
DB 5711 AGATTCAAGCTAGCCAGATGCAACAGATGAAGAAAAACAAGCAGCTGATGCGAAGCA 5770  
QY 3902 AAGAAAGTGAAGCAAGAAAAAGTTATCTATTGGGAACCTTGCTTCTATCACTTAG 3961  
DB 5771 ATACTGAAATGTGTAAGCAATCAAGCCATTTCAACAGCAACTACTTAACGCAAGTTG 5830  
QY 3962 TGAACATGAAAGTTCTGCAAAATCAGAAAAATATACAGTAGCAGAGAACTGAAAGCC 4021  
DB 5831 ATGAACTAAAGCAATGCGAAGCAGCACTTAATGCGTAACACAAAAGTTGTGAAGA 5890  
QY 4022 AAAAAATGATGTTGATGTCACCTGCTTATCAAGCGGACACCAAGTGACAGAGCTTTAA 4081  
DB 5891 AACACGCGGCTAAAGATGAATGATCAATTACAGCAACGCAACAAATGTTATCAATTA 5950  
QY 4082 ATTACAAAGCTGAAGATCAATGGAAGTGTAGGGCTACTGATGCTGTTGCCAATTTAA 4141

DB 5951 ATGATCAGACGCTACACAGAGAAAAAGAGCAGTATTCAACATTTAGCAACAGCAG 6010  
QY 4142 ACAACAAGTAAATAGCTTCTATTAGTGGAGATTAATTAAGTTATGAGGGGAGC 4201  
DB 6011 TTACAGACGCCAAAAATATATATACAGCTGCACATGATGATTAATGTTAGTCAGGCGA 6070  
QY 4202 CAAGAAGCTTTTTCAGCAACCTCAAGTACGCTGCAGTGCAGCGGAGGAGCAATTA 4261  
DB 6071 AAGACGCTGGAAGAAATTCATTTCAAGCAGCAACGCAACGACGCGTTAAATCAATG 6130  
QY 4262 GTTCTGAGCGGSAATTAGAAATTTATCAAGGGGCTGTTCTGTCATTAAGATTGACATG 4321  
DB 6131 CTAATAAATGATGTTGATCAAGCTGACACACTCAAAATCAAGCAATGATATATACACTG 6190  
QY 4322 ACGTGAAGCTAGCGTTGATTAATTTCCATGCAAGAGAGCTTAATGAAATCAATGTCATTG 4381  
DB 6191 GTGCTACAACTGAAGAGAAAAATGCAGCAAAAAGATTTGTTTAAAGCTAAAGAAAAAG 6250  
QY 4382 CCAAGATGTCAAGGAAGTTCTGATCTAGCAAAAAGAAATATCAGGCTTTACTAAATGGA 4441  
DB 6251 GGTATCAAGATATCTTAATATGCAACAACACTTAATGATGTTAGCAAAATTAAGATCAG 6310  
QY 4442 AAGATAAAAAATTTAGAAAGTCTGTTATTAATATACACTGGAATGTTATTAATACGA 4501  
DB 6311 CAGTTGCTGATATTCAAGGTATTACTGCAGATACAACTTAATTAAGATGTTG-----CGA 6364  
QY 4502 AGGAACAATGAAAAAGCAAAAGAAAAAGAGAGCGGCTCATTTGTAATGCTGCTTAT 4561  
DB 6365 AAGATGAATTGCAACAAAGCAAAAGCAAAAGAGCGCTTATTCACAACTGCAGATG 6424  
QY 4562 CGGTTGCTGGAAGGATTAATCCGCTGAGAGGTAGCTATTGCACTCAATCTGTTAAA 4621  
DB 6425 CGACTCTGAAGAAAAAGAACAGCAAAATCAACAGTAGACGACCAATTAACACAAGTA 6484  
QY 4622 ATAAATTTAAAGCAAAATTTGAGTGAAGCAATTAAGAAACCGGAGAGATTAATTAATCATG 4681  
DB 6485 ATCAAAATTAATGAATAATGACAGCTGATGATGATGAACATGCAAAAGATTAATGCA 6544  
QY 4682 CGAAACATGTAATGAGAGCAAAATCATCTACTGTTGTTGTAATGCGCTTTCGAC 4741  
DB 6545 TTCACGAATTAACCAATTCACAGCATCAACAGATGTTAAACGAATGCAGAGC-GGA 6603  
QY 4742 TTGCTATCAGCAAAATGCTTTTTCAGGAATGGAATTCGACATGCGACATGCGAAGCTTATCA 4801  
DB 6604 TTGCTAATCGAATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6657  
QY 4802 ATGACACGATTCGAAAGGTGATTAAGAAAGAAATTTGCTGATTCCTTAATGTCAGC 4861  
DB 6658 -----ACTAATGAAGAAAAAGTTAACGATTTTGACACAGTTAGACGATATGAAGAG 6712  
QY 4862 CAATTAATTCATTTCTGGGAGTAATGTTGCGGAACCATTCGCGTTCTTTCTACGG 4921  
DB 6713 GTTAAATTAATTAATTAATGACCACTACTACAGGATGATTAATTAATTAATTAATTA 6772  
QY 4922 CGGTAGAGCTCTTTTGGAATTAATCTCTTCAATTAATTAATTAATTAATTAATTAATTA 4981  
DB 6773 CAGTACAAAAAGTTCAACCACTTCATGCAAAATCTGTTAAGAAACGACGAGTTAAAAAG 6832  
QY 4982 GAACGAAGTAATCTTTTATGTAAGAAATATCAAAAGTCAATGATACAGCTTTGAATG 5041  
DB 6833 AATTGATTAAGCTGACGTGATTAAGAAACCAAAATTAAGAAACCAAAATGATCAATC 6892  
QY 5042 ATTTCTATTAACAAGCTTTGCTGAGCGCTGCAAGTATTAAGCAGGCTGCAATG 5101  
DB 6893 AACCAAGAAATTAATGATGCAAAACAAAGAGTTGATCTGAATTAATCAAGCGAAGCA 6952  
QY 5102 GAGGAATGATATCTGCAATCGGTTCTGATGAAGAGGAGCTTATGATGATGATG 5161  
DB 6953 ATGTGATCAATCAATCAAAATGAATGATGATGATGATGATGATGATGATGATGATG 7012  
QY 5162 AGTTGAAGAGTAAGTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5221

Db 7013 AAATTAATGCACTTAAACATTAGTGATGACAAAAAAGATGCTTTAGCTAAATTTGAG 7072  
Qy 5222 TTGCCGGAATSCAATGAGAGAAAAGCGCTGAGTTGGAGCAACAGTTGCTCATCA 5281  
Db 7073 ATGCATATATATCTTAAAGTAAAGCGAGATTAACCTTACGCAATGCACTTAAAGTAA 7132  
Qy 5282 ATATTGAAAAAACAATCAGTTATAGCTATTTGTAATAAACAAGTAAATTAACAAGGAAATG 5341  
Db 7133 TTGCTGAAGCGAAACAAAACTTGCTGAATTAACAACTCGGATCAAAATGTTAATC 7192  
Qy 5342 ATCAAGATAGAAAAATATCAATGTGACCTGCAAAAAGTTATATGACCAATCTATATG 5401  
Db 7193 AAGCTACTTCTAAAGATGACATTTGAAGTTCAATTTCAATATGACTTAATATATTAACG 7252  
Qy 5402 CAGTCGAGTTGGAGAGCAAAAGAGCGCTGTGCAAGAGAGCTTCTGCAAGTACTACCT 5461  
Db 7253 ATTACACATTTCCACAGGTAAAGAAAAGATCAGCTACACAGATTTATATGCTTATGCAAG 7312  
Qy 5462 TGAATTAAGACAGTTCTCTCATGTTGATCAAACTGATATTTGACAAAGATTTAGAGAG 5521  
Db 7313 ATCAGAGAAAAATATATATTTTCAAGCTGACATTAATGCAACACAGATGAAAAAGCAAG 7372  
Qy 5522 AAAATTAAGAAATAAGAAAAAGCAAAATGTTATGTTCTAGCTAAATTAACAGTCAAG 5581  
Db 7373 CAATTAAGCAAGTTGACCAAAATG---TTCAAACTGCAATTAAGAAACATTAATTAATGCTG 7429  
Qy 5582 TGGTCACAAAATGCGACAGTCTTCCGAGCAAGTGAAGCAAGCTGAGAGAGCTGAG 5641  
Db 7430 TGGATATATGATGATGCTTATGATGATTAACCAAGTTAAAGCAAGATTTGCTATATTC 7489  
Qy 5642 TAGCAGTTAATTAATTAACAAATACTTCTGCAATATTAATAATAGTACTCAAAATG 5701  
Db 7490 AAGTATGATGCTACTGTTAACTTAAGCGAACCAAGCTATTTAGATTAAGCAGAAATG 7549  
Qy 5702 TACGAAATGCTTTGTTAAAAAG-----CAATCTCATTTATCTTA 5740  
Db 7550 CGAAAGAAATCTATATGATCAAAAGTGACCAAGTTAAGCTGGAAGAAAAAATGAAAGATTAAG 7609  
Qy 5741 TTAATAACAATTGGAATTTGAGCTGAGAGTTGAGAGCTGAGAGAGCTGAGAGCTGAG 5800  
Db 7610 CAATATTAATTAACAAATTTACAGATCAAGCTAAACAAAGTATTAATGATGACACACAACTG 7669  
Qy 5801 TAGCAGTAATTAAGATTGTAATTAATACGATAGCAAGAAATTAATCATGCAAAATCACTG 5860  
Db 7670 CTGAAGTTGAAAAAGCAAAAGCTCAAGAGCTGGAAGCAATTTGATTAACATTCAAATCGACT 7729  
Qy 5861 CGAAGGAAATGTCGGAATTTATTAAGAGCTGATGCGGTAAATTTGCTATTTATGACAGAA 5920  
Db 7730 CAACAGAAAAACAAAAAGCTATCGAAGATTAAGAAATGCACTAGACCAAGATTTGAAGCAG 7789  
Qy 5921 CAGTGTGAGAGTCCGCGTGCAGCAATAGAGAGCTCAACAGTGTGATGAATTTACAG 5980  
Db 7790 GTGTAAATGTCAACGCTGATGCTAACAATGAAAGAAAAAGAGCTTTACGAATGCTTTAG 7849  
Qy 5981 GATTTACAAAAGCAATATGTAAGATTTCTACAGTATTTGCTAAAGAGAAACAGATGAT 6040  
Db 7850 AAGACATTTTATCAAAAAGCACTGAAGATTTTCTGATCAAACTCAATATGACAGAAATCG 7909  
Qy 6041 ATATTACTACTCAAGGCAAGTAAAGTAAAGTGTGATTAAGTATTTCAAAAATCTTAATA 6100  
Db 7910 CTACTGTCAAAAATATGTCGCTTGAACAACTTTAAAGCAACAGTATTAATCTCTGAAGTTA 7969  
Qy 6101 TTAACGAAGACTTATCACAAAAAAGAAAAATAAGTATTAATAAAGATTTGTTACCAATA 6160  
Db 7970 AGAAAAATGCTTTGGAAGCAATCAAGAAATGCTTTAACAAGCAAAATAGAAATTAATTA 8029  
Qy 6161 GTTACGACTACTACTT 6178  
Db 8030 ATGACAGATGCAATGCAAT 8047

; Sequence 1280, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIORITY FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIORITY FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1280  
; LENGTH: 1039  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1039)  
; OTHER INFORMATION: unsure at all n locations  
US-09-902-540-1280

Query Match 0.8%; Score 76; DB 3; Length 1039;  
Best Local Similarity 44.1%; Pred. No. 7.8e-06;  
Matches 403; Conservative 0; Mismatches 507; Indels 4; Gaps 2;

Qy 144 AAAAATGAAAAAAGGATTAATGTTATGACATTACTACAAACAGATTCAGAGGAGAA 203  
Db 107 AAAAAACAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 166  
Qy 204 CGCTTTAAGCTTTTATATGATTTGCTTAAACAGAAATAATATAGCAATCTATATTT 263  
Db 167 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 226  
Qy 264 TGGGAAAAAGATGATCGGAGGTAATTAATCTTTTAACTTTGTCATGGAATAATGA 323  
Db 227 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 286  
Qy 324 AGTATGAGGATTTCAACGGAATTCGAGAAATTAATAATGAGAAATTTATATTTCTT 383  
Db 287 AAAAAATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 346  
Qy 384 AAGCTCGGAGAGATGCGCAGTAGGAAAAAATGAGATTATCAATGCTGTTCTTTTCAATC 443  
Db 347 AAAAAATTAATTAATTAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 406  
Qy 444 TATTATTCAAAACAGATGATTTTAAGAGAGCTTTGGAAGAGCAACATGATTAAGT 503  
Db 407 AAAAAAAGAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 466  
Qy 504 TTTTAATGGAATCATTCAGATGAGAAAAAGTAAAAATTCATTGAATCCGAATGGAAG 563  
Db 467 CCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 526  
Qy 564 CATTACGTTAGAGAAAAATCAATGCTGTTGAAGGCAATGCTTTATATGAGGAGATAT 623  
Db 527 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 585  
Qy 624 TAGATTTGAAGATCTGCAATATCTAAGACAGAAATTAACAGATTTTAATAATTTAGTCA 683  
Db 586 AAAAAATCAATTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 645  
Qy 684 TATTATGATGGAATTAATTTGCTGTCGACCGAGATTTAAAGCTTACCAAGCAAAATC 743  
Db 646 CTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 705  
Qy 744 TGGAGATTAATTTCTTTCAGCTCAATGATTTCTCCCAAAAGC---TATGGGAAAAAA 800  
Db 706 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 765  
Qy 801 TTCAACTGTTGGAAGAAATAGAAAGATATGTAAGAAATAACAAAGCAAAATATTGA 860



QY 7344 AGCAGATATTTCTACAGAGAACTCTTGTAAGCGGATAGAGATACTCAGGAGA 7403  
DB 1438 AGAATGAAAAAGAGAGTGCCTACCACTTAATAGAACTGTGGAACAGGCGAAGAGA 1497  
QY 7404 AGGAATTAAGCCATTGTGGAAGACTTCTGATATTAATGAAAAAATGTGGA-----TAT 7457  
DB 1498 GAGCAAGATCAAAATTAACGAAATATTGAAAAATTAGAAAGAAATGCGTGAAGTAA 1557  
QY 7458 TACAACAGAGACAAGAAATATATATCTCTCTGCTGTGTTGGGAATGCGAGTCTTGC 7517  
DB 1558 TGAAAAAGTTGCAGAGAAATTTAGAGAAATTTAAACAACTGTATTTAATCTGTATGGA 1617  
QY 7518 TTCCCATCAGAGACAGTGCAGTTACAAATTAATAAGAAATTCGAGTTACTGTTGA 7577  
DB 1618 TAAAGTAGAGAAAACGTGAATATTAGCGAGAAAGTTTGAAGAAACATGAATGATTA 1677  
QY 7578 AAATCTTTTGTGAAGACGCTGAAAAAGTAAATGTTAGATCGATATTACGGAATGT 7637  
DB 1678 ACCATTTTGTAGAAATTTGATTAATGTTAAAGGAATACAGAAATTTATTAAACAG 1737  
QY 7638 TGCTTT 7643  
DB 1738 TATGTT 1743

RESULT 15  
5231168-1  
; Patent No. 5231168  
; APPLICANT: DZIEGIEL, MORITEN, BORRE, MARTIN, JERSEN, SOREN,  
; VUUST, JENS, RINECK, KLAUS, WIND, ANNETTE, JAKOBSEN, PALIE H.  
; TITLE OF INVENTION: MALARIA ANTIGEN  
; NUMBER OF SEQUENCES: 19  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/409,658  
; FILING DATE: 18-SEP-1989  
; SEQ ID NO: 1:  
; LENGTH: 3095  
5231168-1

Query Match 0.7%; Score 64.6; DB 9; Length 3095;  
Best Local Similarity 42.0%; Pred. No. 0.0023;  
Matches 657; Conservative 0; Mismatches 884; Indels 22; Gaps 4;  
QY 8163 ACTAACAGTTCTGATTTGGAGAGTTGGAGTCAACAAGAGTACTGTCAGG 8222  
DB 849 AAAAAAGAAATGATCTGAAGTTGTGAAGAAAAAATAATTCACAAGATCAAGTTGA 908  
QY 8223 TAAAACTATGTAGAAAGTTGAGAAAGAAATTTGTTGAGAACAAATCGATGAATGCAAT 8282  
DB 909 AGAAATTCAGTAAATGAGATGAATTTGAGATGTTCACTGAACAATTAATGATTTAGA 968  
QY 8283 TTCTAAAGTAGAAGTTGGATGAAGATTAAGTAATCTCTGCTGATATCAGG 8342  
DB 969 TATATAACAGTTGATCCAGAAATAGTAGAAGTTGAAGAAATTCCTTCAGAACTACATGA 1028  
QY 8343 AAATGAGAGAGAAATTCGCGAGCAAGAGTGAATCTTCTACAGCAACAAGTAATATGA 8402  
DB 1029 AAATGAAGTGCATCCAGAAATTTGTGAATTTGAGAAAGTTTTCCTGAACCAATCA 1088  
QY 8403 ATCCGTAGTTGTTTACGAAGCAAGATTAATGAATTAATGATTAACAAAAAATATAT 8462  
DB 1089 AATATACGAATTTCAAGAAATTAATGAAGATTAATGAAGTGCACATATTCACATGAAT 1148  
QY 8463 TTCAGAGTCAATGCTCTTGTAAATGATCAAAAGAAATGAAGGAATATGAATCTTT 8522  
DB 1149 AATAGAGTAGAAGAAATCTTCCAGAGATGAATTAATAAGTGAAGAAAGTTGAACATGAAT 1208  
QY 8523 AGCGTAGCCGATGTGATGCAACAAGAAACAACAAGCATTTACAGATCAACAAGTT 8582  
DB 1209 AATTAGAGTTGAAGAAATTTCTACAGAAATATTAATAATG-----AAAAAGTCA 1256  
QY 8583 AACTTCTACACTGTAAATGAGAAACGTATATCTCACTTGTGCAAAAGCTTTGCTTA 8642

DB 1257 ACATGAATTAAGTAGAGTTGAAGAAATTTCTACAGAAAGATATTAATAAGAAAGTTGA 1316  
QY 8643 AAATGAAAAATTAAGAAATGTAAAGAACTGAGAGACCTTAATCGGAGCGAAACAGC 8702  
DB 1317 ACATGAATTAAGTAGAGTTGAAGAAATTTCTACAGAAAGATTAATAAGTGAAGTCAACA 1376  
QY 8703 AGCGTTGAAAAATTAACAAGAGTACTACAGAGCATTTGTTGCAGAGAAATTTGGAAAT 8762  
DB 1377 TGAATATGTAAGTTGAAGAAATTTCTACAGAAAGATTAATAAGTGAAGTGAACATGA 1436  
QY 8763 TGAATATTAATTAAGAAAGTATTCAGAGATTAATACATTTGTAAGATCAACGAGACGG 8822  
DB 1437 AATAGTAGAAGTTGAAGAAATTTCTACAGAAAGATTAATAAGTGAAGTCAACATGAAT 1496  
QY 8823 AACCAAGAGAGTCTGTGCGAAAGAAATGTAATTTCTGTGAAAAATACAAATTCAGGGA 8882  
DB 1497 AGTAGAGTTGAAGAAATTTCTACAGAAAGATTAATAAGTGAAGTCAACATGAATGT 1556  
QY 8883 AACCAATCATCATTTGAAGATTAAGCCAGAAATTTGGAAACCGAAAGTGTAAATGTAGA 8942  
DB 1557 AGAAGTTGAAGAAATTTCTACAGAAAGATTAATAAGTGAAGTCAACATGAATGTAGA 1616  
QY 8943 TGCTTTGAATGAATCTTGATGTAGATCTACAGAGAAAGAAAGTGTGCTATGCTGCAATGG 9002  
DB 1617 GGTGAAGAAATTTCTACAGAAAGATTAATAAGTGAAGTCAACATGAATGTAGA 1676  
QY 9003 TATTGAAATGTTGATGTAAATTAATGTGATTAAGAAATGTAGAAGCCAAATTCGAAAG 9062  
DB 1677 GGTGAAGAAATTTCTACAGAAAGATTAATAAGTGAAGTCAACATGAATGTAGAAGT 1736  
QY 9063 ACATGCTAATTTAGAAATCTAGAAACAAAGAAATATCAAGCATTTACAGAGCAAAAGT 9122  
DB 1737 TGAAGAAATTTCTACAGAAAGATTAATAAGTGAAGTCAACATGAATGTAGAAGTGA 1796  
QY 9123 AAATATTTTGAAGAAAGAGACGCTGCGCTGCGCTGCAATTCGAATGTACATTTTC 9182  
DB 1797 AGAAATTTCTCAGAAATTTGTTGAATTTGAAGAAAGTACCAACAAATTAACATGA 1856  
QY 9183 CAATGAGATGATTTAAATTAATTTGGCAAGAGATGATCATCTTCAATTAATTAACAA 9242  
DB 1857 AAATATTTGAATCTTAATAACAGAAAGAAATTAATTAATTTAGTGTGAAGAAAGC 1916  
QY 9243 AAATTCAAAAATATATATTAATTAATTAAGCATCAAGTATGCAATGTGTTATG 9302  
DB 1917 AATTCACAAGAACCCGATGATACCTTAATTAAGAAATGAAGAAACGTTATCCCAAAAC 1976  
QY 9303 GGTGCTGAGCAAGAGGTGCAAGAGCCAAAGCCAGCTTGTGTAAGAAATCAATTA 9362  
DB 1977 ATCTGAAGGTGAATCCACTTAACAGATATGTTCAATTAATAATGTACAGAAATTA 2036  
QY 9363 TAGAATTAATTAATGTGATTTAGCAGAAATTAATAACAGAGGAAACATCAATGTATA 9422  
DB 2037 ACCAAATTAATAAGAAACCAAGATGTAGATGTCCAAACATGTGAACAATAATCA 2096  
QY 9423 TGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9482  
DB 2097 AGAAGATGATTAATGTAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2150  
QY 9483 CAAAGATATGCTGAGCTGCTTGCGCACTGCGCACTTTGA-AAAAATGAAGTAAAT 9541  
DB 2151 AAGAAAGATATGATGAAGAAATTCATCAATTAATAATTAATAAGAAATATCTTTTATAC 2210  
QY 9542 TTAATTAATGCAATCGAATTTTAAAAATTAATCTGCAAGATTTGAAGAGAAAGCTATA 9601  
DB 2211 ATATATATCTCAAGAAATTTTAAAAAGTATCTCAACATTTGTAGTGTATTAATTA 2270  
QY 9602 AAAAAAGTGTGATGATCTTA-----TCAGGTAGATGTGTATACGATTAATTAATCATG 9658  
DB 2271 TGCAATATATGTGTATTAACAGTTGTGAAGTGAATTAAGAAATGAAGAAATATATGT 2330  
QY 9659 ATTCTTGAAGAAAGCATACAAATTAATGATCAATCAATCAAGAGAGAGAAAGGGA 9718  
DB 2331 AATATTTTCCAAACATTTAATTAATTAACAAAAAATTAATAAATTAATAAATTAATAA 2390

Qy 9719 AAA 9721  
Db 2391 ATA 2393

## RESULT 16

US-09-902-540-1357  
; Sequence 1357, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Hinkle, Barry S.  
; APPLICANT: Goldman, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 1357  
; LENGTH: 612  
; TYPE: DNA  
; ORGANISM: Myxococcus xanthus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(612)  
; OTHER INFORMATION: unsure at all n locations  
US-09-902-540-1357

Query Match 0.7%; Score 63.4; DB 3; Length 612;  
Best Local Similarity 48.0%; Pred. No. 0.0024;  
Matches 178; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

Qy 7110 AAAAAAGTAATGTAACGCAAGAGAAAAATTTAAGCAACAGCAAGCAATGC 7169  
Db 44 AA 103  
Qy 7170 AGCAATCGAGAGAGCAATCGCAAGCAATGCTTGTAAATTAATTTGCAACAGCTGT 7229  
Db 104 AA 163  
Qy 7230 AGAAGATAGAAAAAATTTCTGAAGAAAGAAAGCAAGATTAACTTTAGACGAGT 7289  
Db 164 AA 223  
Qy 7290 TAACAAAGACAGATAAAAAGTAATGATGCTACGAAAAAATTTACATCAGCAGG 7349  
Db 224 AATTAATAA 283  
Qy 7350 TATTTCTAGAGAGATCTCTGTAAAGCGATAGAGAGATCTCAGGAGAGAGAT 7409  
Db 284 ATATTAATAA 343  
Qy 7410 TAAAGCATTTGAGAGATCTCTGATATTAATGAAAAAATGTAGATATTACACAGAGA 7469  
Db 344 TAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 403  
Qy 7470 CAAGATTAATA 7480  
Db 404 AAAAAAAAAATA 414

RESULT 17  
US-09-662-254B-24/C  
; Sequence 24, Application US/09662254B  
; Patent No. 6931145  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Li, Yi  
; APPLICANT: Bawden, Alison Louise

; TITLE OF INVENTION: Materials and Methods for Delivery and Expression of Heterologous  
; FILE REFERENCE: US-221C1XCI  
; CURRENT APPLICATION NUMBER: US/09/662,254B  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 09/086,651  
; PRIOR FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/224,479  
; PRIOR FILING DATE: 2000-08-10  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 24  
; LENGTH: 50000  
; TYPE: DNA  
; ORGANISM: Ambacta moorei entomopoxvirus  
US-09-662-254B-24

Query Match 0.6%; Score 62.4; DB 3; Length 50000;  
Best Local Similarity 43.8%; Pred. No. 0.016;  
Matches 469; Conservative 0; Mismatches 586; Indels 15; Gaps 4;

Qy 10 ATCAAAATTAACGTTCAAGAGACAAAGAGATATCAGATCTTAAAAAGTTTAAAG 69  
Db 40177 ATGGAATATTAATGATTAATAATTAATTAACAAAAATTTAATTAATTAATTAATG 40118  
Qy 70 ATTTGGATTTGATTAACATGACGATGAGGCGTAATGATACATCACCGCACT 129  
Db 40117 GAATCGAAATGTAATAATTAATTAATTAACAAATTTAATTAATTAATTTCCGAAAT 40058  
Qy 130 GAGAAATTTGCAACAAATAGAAAAAGATAATGTTTATGACATTAACAAACAG 189  
Db 40057 AAAAAATATATTAATTAATTAATTAATGATTAATGTTTCAATACATCAATATAT 39998  
Qy 190 ATTCAGGAGAGACGCTTTTAACGTTTAATTAATTTGCTTTAACGAAATTAATA 249  
Db 39997 ATTAATGTCCTCAATGATTAATGAAAAATTAATTAATTAATTAATTAATTAAT 39939  
Qy 250 GCAATCTATATTTTGGGAAAAAGATGAGGAGGTAATTAATCTTTTAATTTGTC 309  
Db 39938 ATTAATCTATTTGTCCTTAACAAATGTAATAAAAAATTAATTTTCGATTTAAGCC 39879  
Qy 310 AATGAAAAATTAAGATGATGATTAATCAACGAAATTCGAG-----AAATTAAT 362  
Db 39878 ATTAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39819  
Qy 363 TCGAGAAATTAATTTCTTAAGCTCGAAGGATGCGATGAGAAAAATGAGTAT 422  
Db 39818 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 39759  
Qy 423 CAATGCTGCTTTTCAATCTATTAATTCAGAAACAGATGATTTAAGAGGCTTTGA 482  
Db 39758 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39699  
Qy 483 AGAAGCAACATGTAAGTTTAAATGAATCAATTCAGTATGAAAAAGTAAAAAT 542  
Db 39698 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39639  
Qy 543 TCCATTAATTCGAATGAGATTAATGATGAGAAAAATCAATGCTTTGAAGCAT 602  
Db 39638 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39579  
Qy 603 CGGTTTATGCGGCGGATTAATGATGTAAGATCTGCAATTAAGCAAGAAATTA 662  
Db 39578 ATTTTCAATGCTTCTGATGATGATTAATTAATTAATTAATTAATTAATTAATTA 39519  
Qy 663 AGATTTTAAATTTAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 722  
Db 39518 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 39459  
Qy 723 AAAAGTACCAAGCAAAATCTGAGATTTATTTCTTCACTCAGATGAT-----TC 776  
Db 39458 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 39399

Qy 777 TCTCAAAAAAGCTATGGGAAAAAATTCACCTGTGGAAAGAGATAGAGATATGTAAA 836  
Db 39398 TCTACATCTCTATTATTATTAATCACTTTATAGATTGATTAATTAATTAATTAATTA 39339  
Qy 837 AAGAAATACCAAGCAAAATATTGAATCTG-ATGCTGTATTGGAACAGATGAAATATTA 895  
Db 39338 AATATATATTTATTTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 39279  
Qy 896 AAATTAGTCGCAAAAGCTACAAATGGAGATTATTAAGAAAGGAAAAAGAACTT 955  
Db 39278 AAATATAGAAATTTATTAATAAATAATTTATTAATAAATAAAGCTATATTA 39219  
Qy 956 ATACACTCTTTAAGTTATTCAGATGTGAAAGCTTCGTAAGATTAATAAGGAAAG 1015  
Db 39218 ATTGACAAATCTATTATTTATCTATCTATTAATGAAAAAACCAGAAATTAATAAG 39159  
Qy 1016 TCATAGAAAGAAATGTTGACATTACAGCTGAAGCAAAAGATTCTATGAT 1065  
Db 39158 AAGTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 39109

## RESULT 18

PCT-US93-07261-10  
; Sequence 10, Application PC/TUS9307261  
; GENERAL INFORMATION:  
; TITLE OF INVENTION: P4EMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREOF  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John H. C. Blaisdale  
; STREET: One Giraldia Farms  
; CITY: Madison  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07940-1000  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: Macintosh 6.0.5  
; SOFTWARE: Microsoft Word 5.1a  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07261  
; FILING DATE: 19930805  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/927,531  
; FILING DATE: 07-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blaisdale, John H. C.  
; REGISTRATION NUMBER: 31,895  
; REFERENCE/DOCKET NUMBER: DX0288K  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-822-7398  
; TELEFAX: 201-822-7039  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4766 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium falciparum  
; STRAIN: Malayan Camp  
; IMMEDIATE SOURCE:  
; CLONE: p2b1:p12-1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 3..4766  
; PCT-US93-07261-10

Query Match 0.6%; Score 62.2; DB 6; Length 4766;  
Best Local Similarity 42.7%; Pred. No. 0.0083;  
Matches 551; Conservative 0; Mismatches 728; Indels 12; Gaps 4;

Qy 8271 ATTGAATGCAATTTCTAAAGTAGAAGGTTGGATGAAATGAATGAATCTGCTAAATCTTC 8330  
Db 2363 ATATTAATGAATCTGATTTAGCAAAAGGTAAGAGTAATTAATTAAGCAATGAAATCT 2422  
Qy 8331 TGTGTATCAGAAAATGAGAGAGAAATTCGCGAGCAGAGATGAATCTTCTACAGCA 8390  
Db 2423 AGAAGAAATTAATGAATCACTGATTTAGCAAAAGGTAAGAGTAATTAATTAAGCAGCGA 2482  
Qy 8391 AAGTAATCTGATTCCTGATGCTTTAGCAAGACAGATTAATGAATTAATGAATTAC 8450  
Db 2483 AAATCTAGAGAAATTAATGAATGAACTGATTTAGCAAAAGGTAAGAGTAATTAAGC 2542  
Qy 8451 AAAAAATATATTTTCAAGAGTCAATGCTCTTGTCTTTAATGATCAAAAGATGAAGCAA 8510  
Db 2543 ACGGAAATCTAGAAAGATTAATTAATGAATGAACTGATTTAGCAAAAGGTAAGAGTAAT 2602  
Qy 8511 TATGAAATCTTTAGCGGTAGCGCGTGTGATGCAAG- - - -GACCAACAAAGCAATT 8565  
Db 2603 TAAAGCAGCGCAAAATCTAGAAAGATTAATGAATGAACTGATTTAGCAAAAGGTAAGAG 2662  
Qy 8566 ACGAGATCAACAGATTAACTTTCAACATGTA- - -ATGAGAGAAACGATCTCAACTCG 8624  
Db 2663 AACTTAATTAAGCAGCGAAATCTAGAAATTAATGAATGAACTGATTTAGCAAAAGGTA 2722  
Qy 8625 TGCAAAAGCTTTGGCTAAAAATGAAAAATTAAGAAATGAAGAACTGAGAGAGCTT 8684  
Db 2723 GGAATGACTTAATTAAGCAGCGAAATCTAGAAATTAATGAATGAACTGATTTAGCAAA 2782  
Qy 8685 AGTCGAGCGGAAACAGACGCGTTGAAATTAATCAAAAGATCTACAGAGCAATGCT 8744  
Db 2783 AGTAAAGAGATTAATTAATTAAGCAGCGAAATCTAGAAATTAATGAAGAAAGATTA 2842  
Qy 8745 TGCAAGAAATTTGGAAATTTGAGATTAATTAATGAAGAAATGCAAGATTAATGATGCT 8804  
Db 2843 TATGAATATATATATTAATTAATTAAGATCTGATGATTTAAAGAAATGCAAGACT 2902  
Qy 8805 AAGAGTCAACGAGACGAAACCAAGAGAGCTTTGTGCGAAAGAAATGATTTCTGTGAA 8864  
Db 2903 AAAAAATTAAGAAATTAAGAAATTAAGATCTGATGATTTAAAGAAATGCAAGACTTAA 2962  
Qy 8865 AAATTAATTTCAAGGGAACAAATCATCTGATTAAGATTAAGCAGAAATGTTGGAAC 8924  
Db 2963 AAATTAAGAAATTAAGAAATTAAGATCTGATGATTTAAAGAAATGCAAGACTTAA 3022  
Qy 8925 CGAAGGTAAATGATGCTTTGAATGAATGAACTGATGATTTCAAGAGAAATGAG 8984  
Db 3023 TAAAGATTTAGAAATTAAGATCTGAGAGATTTAAAGAAATGCAAGATTTAAATTA 3082  
Qy 8985 TGCTATGCTGAAATTTGATTTGAAATGTTGATGTAATATGATTTAAAGAAATGAT 9044  
Db 3083 AGAATTAAGAAATTAAGAGATCTGAAGATTTAAAGAAATGCAAGCTTAAATTAAGA 3142  
Qy 9045 AGAAGCCAAATCGAAGACATGCTATTGTAGAACTGAGAAACAGAAATTCAGAC 9104  
Db 3143 ATTAGAAATTAAGAGATCTGAAGATTTAAAGAAATGCAAGATTTAAAGAAATTT 3202  
Qy 9105 ATTTAACAAGACAAAGTAATATTTCTTGAAGAAAGAGCGCTGACGTGAG- - -CTG 9160  
Db 3203 ACAAATTAAGAGATCTGAAGATTTAAAGAAATGCAAGATTTAAAGAAATTAAGATTA 3262  
Qy 9161 CAATATCGAATGTACACTTTCCATGAGATG- - -GATTTAAATTTGCAAGAGAGTA 9218  
Db 3263 AAATTAAGATCTGAAGATTTAAAGAAATGCAAGAGCAAAATTAAGAAATTAAGATTA 3382  
Qy 9219 TGCAATCTTCAATTAATTAACCAAAATTTCAAAATTAATTAATTTAGCATCAAGTAG 9278  
Db 3323 TAAAGATCTGAAGATTTAAAGAAATGCAAGATTTAAAGATTTAAAGATTTAAAGATTA 3382  
Qy 9279 TGAATCGAATGATGATGCTTGAAGGTCGTAAGAGCAAGAGTCAGAGCAAGCAAGCAG 9338  
Db 3383 AGGATCTGAAGATTTAAAGAAATGCAAGACTTAAAGAAATTAAGAAATTTAAAGG 3442  
Qy 9339 AGTAGTGAAGATCAATTAATTAAGATTAATTAATGATTTAGCAGGAAATTTAA 9398

```
Db      3443  ATCTGATGATTTAAAGAAATGCAAGATTAAATAATTAAGATTACGAATTAAGGATC 3502
Oy      9399  AACAGAGGAAACATCAATGTATATCCGGATATGATTAATAATTATTAATAGTAGAC 9458
Db      3503  TGAAGGATTAAAGAAATGTTATATCAATAATGATTTAAAGAAATTAAGATATTCAAA 3562
Oy      9459  AAATCTAAGGCTATTGCGGATGCCAAAAGTCATGCTGCGCTTCCGCACTGCCAC 9518
Db      3563  TAAAGATTATCTAATTAAGACATGAAAAATTAAGAACTATTAAACAAAGATTTCTAA 3622
Oy      9519  TATTGAAAAAATGAAGTAAATTTAATAT 9549
Db      3623  TAAAGACATGAAGAAATTAAGAACTATTAAAT 3653

RESULT 19
US-09-949-016-12147
; Sequence 12147, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FaSeqSeq for Windows Version 4.0
; SEQ ID NO 12147
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147

Query Match      0.6%; Score 62; DB 3; Length 767677;
Best Local Similarity 46.5%; Pred. No. 0.048;
Matches 200; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Oy      2849  ATTCACTATCTGTAAGAGCTTAAGAAAAAGCTGAATTATTCAGAGAAAAATTATTAATGCG 2908
Db      343436  ATTAATGTAATGCTGTAATTAAGAACGATGCTGTGATGAAGATGATGGTGGATGAAG 343495
Oy      2909  ATGCATTGAACAGACTTTTTCATGTAAGCGGATCTTTTAATGCTGCTCAGGTGGGAATG 2968
Db      343496  ATGATGGGGATGTGTGATGAATGAAGATGATGATGATGAAGATGTTGATGATGAATG 343555
Oy      2969  CAATCAATGAATGGGAAGTTAATGAGAGTATCAGTAAGGCAAGATTCATTGATGATG 3028
Db      343556  AAGATTAATGATGATCAATTAAGATGATGATGATGATGATGATGATGATGATGATG 343615
Oy      3029  ACGAAGCATATTGTAAGAGCTTAATTAATAAATTCCTTTAAACAGTAAGATTAATCTCTG 3088
Db      343616  AAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343675
Oy      3089  TTTGGAATGCTGCCGGTTCAAGCGGGAATCGAAACGAAAAATCGCGCGGTGCGGTGCTG 3148
Db      343676  ATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343735
Oy      3149  TTGCGGTAATGATTAATGATTAATTTCAACAAAGCTTCATTGAAGATTAATGACGAAGAC 3208
Db      343736  ATGAAGATGATGATGATGATTAATTAAGATTAAGATGATGATTAAGATGCGGATGATG 343795
```

```
Oy      3209  AAGTAATATGATGATGATTAAGATGATGAATGAATGAATGAATGAATGATGATGATGATG 3268
Db      343796  ATAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343855
Oy      3269  TAGATGCAAA 3278
Db      343856  GTGGTGAAGA 343865

RESULT 20
US-09-949-016-17361
; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FaSeqSeq for Windows Version 4.0
; SEQ ID NO 17361
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17361

Query Match      0.6%; Score 62; DB 3; Length 767677;
Best Local Similarity 46.5%; Pred. No. 0.048;
Matches 200; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Oy      2849  ATTCACTATCTGTAAGAGCTTAAGAAAAAGCTGAATTATTCAGAGAAAAATTATTAATGCG 2908
Db      343436  ATTAATGTAATGCTGTAATTAAGAACGATGCTGTGATGAAGATGATGGTGGATGAAG 343495
Oy      2909  ATGCATTGAACAGACTTTTTCATGTAAGCGGATCTTTTAATGCTGCTCAGGTGGGAATG 2968
Db      343496  ATGATGGGATGTGTGATGAATGAAGATGATGATGATGATGAAGATGTTGATGATGAATG 343555
Oy      2969  CAATCAATGAATGGGAAGTTAATGAGAGTATCAGTAAGGCAAGATTCATTGATGATG 3028
Db      343556  AAGATTAATGATGATCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 343615
Oy      3029  ACGAAGCATATTGTAAGAGCTTAATTAATAAATTCCTTTAAACAGTAAGATTAATCTCTG 3088
Db      343616  AAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343675
Oy      3089  TTTGGAATGCTGCCGGTTCAAGCGGGAATCGAAACGAAAAATCGCGCGGTGCGGTGCTG 3148
Db      343676  ATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343735
Oy      3149  TTGCGGTAATGATTAATGATTAATTTCAACAAAGCTTCATTGAAGATTAATGACGAAGAC 3208
Db      343736  ATGAAGATGATGATGATGATTAATTAAGATTAAGATGATGATTAAGATGCGGATGATG 343795
Oy      3209  AAGTAATATGATTAAGATTAAGATGATGAAGTAACTGAGTAATGCGGGAATCTTTGAAG 3268
Db      343796  ATGATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 343855
Oy      3269  TAGATGCAAA 3278
Db      343856  GTGGTGAAGA 343865
```

RESULT 21

US-08-072-610-1

Sequence 1, Application US/08072610

Patent No. 5532133

GENERAL INFORMATION:

APPLICANT: Barnwell, John

TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,

TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby and Darby

STREET: 805 Third Ave.

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022-7513

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/072,610

FILING DATE: 19930602

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gogoris, Adda

REGISTRATION NUMBER: 29,714

REFERENCE/DOCKET NUMBER: 5986/07686

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)527-7700

TELEFAX: (212)753-6237

TELEX: 236687

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3337 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium vivax

IMMEDIATE SOURCE:

CLONE: PMB3.3.1

US-08-072-610-1

Query Match 0.6%; Score 61; DB 2; Length 3337;  
Best Local Similarity 42.3%; Pred. No. 0.013;  
Matches 456; Conservative 0; Mismatches 615; Indels 6; Gaps 2;

QY 5692 ACTCAAAATGTACGAATGCTTTGTAAAAAGCAAAATCTCATCTATTAAAACAATT 5751  
DB 504 ACAGAGAAGAGACTACCTGTGATATGAGAGAGAAAGAGCTGGCGATATGGAACA 563  
QY 5752 GGAATTGAGCTGAGCTTGAGCTGAGAGAGAGCTGAGAGACAGGTTCTGTACAGTGAAT 5811  
DB 564 GGGGAAGAGCTGGTGAATTTGGAACAGGGGAAAGAACTGGCGATTGGAAGCAGGGGAA 623  
QY 5812 AAGATTGTAATTAATACATAGCAGCAATTAATCAATGCAAAAAATCACTGCGAAGGAAT 5871  
DB 624 GAACTGGCGATTGGAAGAGAGAGAGAGAGAGCTGTGATTTGGAAGCAGGGGAGAACT 683  
QY 5872 GTCGAGTTAATTACAGAGCTGTGATCGGTAAATTTGCTAAATTATGACAGAAACAGTGTCTGA 5931  
DB 684 GCGGATTGGAAGCAGGGGAGAGAACTGGAAGATGCGGAACTGGAAGAGAGAGCAACTGGA 743  
QY 5932 GTGGCCCGTGCAGCAATAGAGAGCTCAACACAGTGA---TGAATTACAGAGATCTACA 5988  
DB 744 GATGCGAAACTGAAATGAGCAACTGTGTATGTAGACAGAGAGATGTTCAAGCTGAT 803

QY 5989 AAAGCATATGTAAGAATTCTACAGATGTTCTAAAGAGAAAGAGATGATTATTAAT 6048  
DB 804 GGAGCAGAAAAAGTACATGTTCTGCTCAGAAAAATGTACAACTGCCGATAGTAATAT 863  
QY 6049 ACTCAAGGCAAGTAAAGTGTAGATTAAGTAAATTCAAAAATCTTAATATTACGAA 6108  
DB 864 GCCCTTTGGAAAGTATTTTGGATTAAGATTAATTTTGTATCATATTAAAGATTTGAG 923  
QY 6109 GACTTATCAAAAAAAGAAAAATAGTAATTAATAAAGATTGTTACCAATAGTTCACT 6168  
DB 924 CCCTATTTCGAACAATTTGTGGCGGTACTGTAAACATGTTACGGGCAAGAAATTGCA 983  
QY 6169 ACTCATCTTTAAATCTTTTATTGGCAAAATCCCGTGTTCAGAGCAAGCCGAG---TG 6225  
DB 984 ATGAACTGTACCATTAACAGTGGCAGAAAGCCCGGCAAGTACCGAGGAAATTA 1043  
QY 6226 GCAGGAAGCTGTAATATCAACAGAGTTATGAGAAACAGAAAGCTCTTTGAAAAATCT 6285  
DB 1044 GATGCCACTCCAGAGATGATCTCGCATTAATGATTACAGAAATCTCCGAGGAAATGAA 1103  
QY 6286 ATATTAATGCAAAACATTAATCTGTAAATCAAGAGATTACAGAAATTCATGGAGTA 6345  
DB 1104 TTAGTATTAGTGAAGAGGCACTGAAAGAAATCAACGAAAGTGGACCAACGAGAA 1163  
QY 6346 GTAGTCTCTGTGTGTGTGTGAATGTAGAGAGTGAAGCTTCTGTATACCAATATT 6405  
DB 1164 GAGCAACCGAAGAAATTAAGATGCCATCCAGAGATGATTTGCCATTAAGCAAACTGC 1223  
QY 6406 ATAAAAAGAAATACCAAGACAAGAGTTGAAAAATCAATGCTGATGAAGTTTCGA 6465  
DB 1224 AGAAGAGAAACAGAAAGAAAGTGAAGGAGAAAGAAACAGAAAGAGTGCAGAAAGAA 1283  
QY 6466 GAAGAAGTGAATTAACAGCAGATTTCTAAGCAAGAAATTCCTTTTGGAGTGGAGTC 6525  
DB 1284 GTATCAGAGAACTCCAGAAAGAGAAAGAGATTAAGGAACTCCAGAGATGATTTTC 1343  
QY 6526 GCACAGCCGGGTGAGAGCCGAGGTGAGCAAGAACGTTTCCGTAAATCAATTTGCAAGA 6585  
DB 1344 GCATTAAATGAACTACTTAATTAAGAAAGAAACGAAAGAACTGAGAGAGAGAAACCTA 1403  
QY 6586 AAGACGAAAGTATGATGTGGAAGAAAGCAAGATTTTGTAAAAAAGCTGAGATTACAGA 6645  
DB 1404 GAGGAGAGAAAGAAACGTAGAGAGAGAAAGAAACGTAGAGGAGAAAGAGTGCAGAAAGA 1463  
QY 6646 AAAGCTTATAGTTCTGTTGCAATTTGGAATGCCAGAGTGGAGTGCCTGCAAAAGAGCT 6705  
DB 1464 GAAGAAGGTTAGAGGCAACTCCAGAGATGACTTCCAAATTTAGAAAGCAATTCAGAGAA 1523  
QY 6706 GGAATTGAGCAGCAGTGGCAGTTACCAAGATGATCAACAGAGCAAGAGTG 6762  
DB 1524 GGAAGAGGAG 1580

RESULT 22

US-08-719-822B-1

Sequence 1, Application US/08719822B

Patent No. 5874527

GENERAL INFORMATION:

APPLICANT: Barnwell, John

TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Darby and Darby

STREET: 805 Third Ave.

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10022-7513

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/719,822B  
FILING DATE: 09/30/96  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gogoris, Adda  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17686US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3337 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium vivax  
IMMEDIATE SOURCE:  
CLONE: PMB3.3.1  
US-08-719-822B-1

Query Match 0.6%; Score 61; DB 2; Length 3337;  
Best Local Similarity 42.3%; Pred. No. 0.013;  
Matches 456; Conservative 0; Mismatches 615; Indels 6; Gaps 2;

QY 5692 ACTCAAAATGTCGAAATGCTTTGTAAGAAAGCAATCTCATCTTAATAACATTT 5751  
DB 504 ACAGAGAAAGAGACTACCTGCTGATATGAGAGAGAAAGAAAGCTGCCATATGAGAGCA 563  
QY 5752 GGAATTGGAGCTGAGCTTGGAGCTGAGAGCTGAGAGCTGACAGCTTCTGTAGAGTGAAT 5811  
DB 564 GGGAGAAAGCTGCTGATATTTGAGAGAGAGGAGAAAGCTGGCATTTGAGAGAGGAGAA 623  
QY 5812 AAGATTGTAATTAATACGATAGCAGAAATTAATCATGCAAAAATCACTCGAAGGAAAT 5871  
DB 624 GAAACTGGGATTTGGAAGCAGGAGGAGAAAGCTGCTGATTTGGAACAGAGGAGAAAGAACT 683  
QY 5872 GTTCGAGTTATTAACAGAGCTGATGCGGTAATTTGTAATTAATGAGAAACAGTCTGCA 5931  
DB 684 GGGGATTTGAGAGCAGGAGGAGAAAGCTGATGATGAGACACAGAAAGATGTTCACTGAT 743  
QY 5932 GTGGCCCTGCGACAAATAGAGCTCAACCACTGTGA--TGAATTACAGATTTACA 5988  
DB 744 GATGCGGAAAGCTGAAATGAGCACTGTGTATGAGACACAGAAAGATGTTCACTGAT 803  
QY 5989 AAGCATATGTAAAGATTCTACAGTATTTCTAAAGAGAAACAGATGATTAATTAATCT 6048  
DB 804 GGGAGCAAAAAGTACATCTTCTGCTCAAGAAAAGTACACCTCCGATAGTAATGAT 863  
QY 6049 ACTCAAGGCAAGTAGATAAGTGTAGATAAGTAAATTCAAAATCTTAATTAATTAACGA 6108  
DB 864 GCCCTCTTTGGAAGTATTTTGAATAAGATATAATTTTGTATCATATTAAGATTTTCGAG 923  
QY 6109 GACTTATCAAAAAAAGAAAAAATAGTAATTAATTAATTAATTAATTAATTAATTAATTA 6168  
DB 924 CCACATTTGCAAAATTTGCGGGGTACTGCTAAACATGTTACGGGACAAAGAAATGCCA 983  
QY 6169 ACTCATACTTTAAATCTTATTTGGCAAAATGCGCTGTTTCAAGACAAAGCCGAG--TG 6225  
DB 984 ATGAAACCTGTACATTTACAGTGGCAGAAAGACCCCGCAATACCAAGGAGAAAGATTA 1043  
QY 6226 GCAGGAAGCTGTTAATATCAACAAAGTTTATGAGAAACAGAACTCTTTGTAAGAAATTTCT 6285  
DB 1044 GATGCCACTTCAGAGAGTACTTCGATTTAGATGTTTACAGAAATCTCCCGAGAAAGTAGAA 1103  
QY 6286 ATATTAAAGCAAAACATTAATTTCTGTAATAATCAGAGATTTACAGAAATTCATCGGAGTA 6345

DB 1104 TTAGATTAAGATGAAGAGCAACTGAGAGAAATCAACGAAAGTGGACCAACGAGAGAA 1163  
QY 6346 GTAGTTCTTGTGTTGGTGTGAGAAATGTAGAGTAAAGCTTCTTGTGATACCAATATT 6405  
DB 1164 GAGCCAAACGAAAGATTGATGCGCATCCAGAGGTGATTTTGCATTAGACGAAACTGC 1223  
QY 6406 ATAAAAAGAAATTAACAGACAAAGTGTGAAAAAATCAATATGTGATGAAGTTTCGA 6465  
DB 1224 AGAAGAGAAACAGAAAGAAAGTAGAGGAGAAAGAAACAGAAAGCTGCGAAGAGAA 1283  
QY 6466 GAAAGAGCTGAATTAACAGCAGATTTCTAAGCAAGAAATTTCTTTTGAAGTCGAGTC 6525  
DB 1284 GTATCAGAAAGAACTCCAGAAAGAGAAAGAGTAAAGGCAATCCAGAGATGATTTTC 1343  
QY 6526 GCAGCAGCCGGGATGAGAGCCGAGAGTGCAGAAACCTTTCCGTAAATCAATTTTCAGA 6585  
DB 1344 GCATTAGATGGAACCTTATGAAAGAAACCGAAGAACTGCAGAAAGAGAAACCGTA 1403  
QY 6586 AAGACGAAAGTATGATGGAAGAGCAAAAGATTTTGTAAAAAAAGCTGAGATTACAGA 6645  
DB 1404 GAGGAGAAAGAAACCTGAGAGGAGAGAAACCTGAGGAGAGAAAGCTGCAAGAGCA 1463  
QY 6646 AAACGTTATGTTCTGTTGCAATTTGAAATGCGCAGTCCGAGTGCCTGCAAAAGAGCT 6705  
DB 1464 GAAAGAAAGTTAGAGAGCAACTCCAGAGATGACTTCCAATTAGAAACCATCAGAGAA 1523  
QY 6706 GGAATTTGAGCAGCAGTGGCAGTTTCAAAAGATGAATCAACAGAGCAAGCAAGATG 6762  
DB 1524 GGAGAGGGGAAAG 1580

RESULT 23  
US-09-092-458-1  
Sequence 1, Application US/09092458  
Patent No. 6231861  
GENERAL INFORMATION:  
APPLICANT: Barnwell, John  
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens, Monoclonal Antibodies, and Diagnostic Assays  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby and Darby  
STREET: 805 Third Ave.  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/092,458  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/719,821  
FILING DATE: 09/30/96  
ATTORNEY/AGENT INFORMATION:  
NAME: Gogoris, Adda  
REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17686US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237  
TELEX: 236687  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3337 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double

TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Plasmodium vivax  
IMMEDIATE SOURCE:  
CLONE: FVMB3.3.1  
US-09-092-458-1

Query Match 0.6%; Score 61; DB 3; Length 3337;  
Best Local Similarity 42.3%; Pred. No. 0.013;  
Matches 456; Conservative 0; Mismatches 615; Indels 6; Gaps 2;

QY 5692 ACTCAAAATGTCAGAAATGCTTTGGTAAAGCAAAATCTCATTCATTAAACAATT 5751  
DB 504 ACAGAGAAGAGACTACCTGGTATATGAGAGAGAGAGAAAGCTGGCGATATGGAACCA 563  
QY 5752 GGAATTGAGCTGAGATTGAGAGCTGAGAGAGCTGAGATGACAGCTTCTGTACAGTGAAT 5811  
DB 564 GGGGAGAAAGCTGGTGAATTGGAGACAGGGAGAAAGAACTGGCGATTGGAGACAGGGAA 623  
QY 5812 AAGATTGTAATTAATACATACATACCAATTAATTCATGCAAAAATCACTGCCAAGGAAT 5871  
DB 624 GAAACTGGCGATTGGAGAGCAGGGAGAAAGCTGGTGAATTGGAGACAGGGAGAAAGAACT 683  
QY 5872 GTCGAGATTATACAGAGCTGTATGCGTAATTGCTAATTATGACAGAAAGTGTCTGGA 5931  
DB 684 GGGCGATTGGAGAGCAGGGAGAAAGCTGGTGAATTGGAGACAGGGAGAAAGAACT 743  
QY 5932 GTGGCCCGTGACAGCAATAGAGCTCCAAACCACTGTGA--TGAATTACAGAGATCTACA 5988  
DB 744 GATGCGGAAATCGAAATAGAGCAACTGTATGTAGACACAGAAAGATAGTTCACTGAT 803  
QY 5989 AAAGCATATGTAAGATTTCTACAGTATGCTAAGAAAGAAACAGATGATATTAAT 6048  
DB 804 GAGAGCAAGAAAGTACATGTTCTCTGCTCAGAAATGTCACCTGGCATATGATAT 863  
QY 6049 ACTCAAGGCAAGTATGATTAAGTGTATTAAGTATTCAAAATCTTAATTTAAGCA 6108  
DB 864 GCCCTCTTGGAGATTTTGGATTAAGATATTAATTTTGGATCATATTTAAGATTTCCAG 923  
QY 6109 GACTTATCAAAAAAGAAAAATTAAGTAATAAAAAAGATTGTTACCAATAGTTCACT 6168  
DB 924 CCACATTCGAAACAATTTGTGCGGGTACTGCTAAACATGTTACGGGACAAAGATTGCCA 983  
QY 6169 ACTCATCTTTAAATTTTATTTGCGAAATGCCGCTGTGTACAGACCAAGCCGAG--TG 6225  
DB 984 ATGAACCTGTACCATTAACAGTGCAGAGAGCCCGGCAAGTACAGCGGAGAAATTA 1043  
QY 6226 GCAGGAACCTGTAATATCAACAAGTTTATGAGAGAAACAGAGCTCTTGTAGAAAATCT 6285  
DB 1044 GATGCCATCTCCAGAGATGACTTCCCATTTAGATGTTACGAATCTCCGAGAGAAATGA 1103  
QY 6286 ATATTAAATGCAAAACATTATTTCTGTAAATCAGAGATTACACGAATTTCAATCGAGTA 6345  
DB 1104 TTAGATTATGATGAGAGGCACTGAGAGAAATCAACGAGAGTGGACCAACGAGAA 1163  
QY 6346 GTAGGTTCTGTGTGTGTGTGTGTAATGTAGAGATGAGAGCTTCTTGTATACCAATTT 6405  
DB 1164 GAGCCAAACCGAAGAAATTAGATCCCACTCCAGAGATGAGTTTCGATTAGACGAATACTGC 1223  
QY 6406 ATAAAAAGAAATACCAAGACAGAGTTGAGAAAACTACAAATGCTGATGAAAGTTTCGA 6465  
DB 1224 AGAAGAGAAACAGAAAGAACTAGAGGAGAGAAAGAAAGAGAGCTGCGAGAGAGAA 1283  
QY 6466 GAAAGAGCTGAAATTTACAGCAGATTCTAAGCAAGAAATTTCTCTTTTGGATCGAGATC 6525  
DB 1284 GTATCAGAAAGAACTCCAGAGAGAGAGAGAGATTAGAGGCAATCCAGAGAGATGATTC 1343  
QY 6526 GCAGAGCCGGGTAGAGCCGAGATGGCAGAGAACCGTTTCCGTAAATCAATTTGACAGA 6585  
DB 1344 GCATTAGATGAACTACATTAGAGAAACCGAAGAACTGCGAGAGAGAGAAACCGTA 1403

QY 6586 AAGACGAGTAGATGTGGAAGAGCAAAAGATTTTGTAAAAAAGCTGAGATTACAGCA 6645  
DB 1404 GAGGAGAGAGAAACCGTAGAGAGAGAGAAACCGTAGAGAGAGAGAAAGCTGCGAGAA 1463  
QY 6646 AAAGCTTATAGTTCTTGTTCATTTGAAATCCCGAGTGGAGTGGCTGCAAAAGAGCT 6705  
DB 1464 GAAAGAGATTAGAGGCAACTCCAGAGATATCTTCCATTTAGAGAAACCATTCAGAGAA 1523  
QY 6706 GAAATTGAGCAGCAGTGGCAGTTACCAAGATGATCAACACGAGAGCAAGAGTG 6762  
DB 1524 GAG 1580

RESULT 24  
US-08-719-821C-1  
Sequence 1, Application US/08719821C  
Patent No. 6706872

GENERAL INFORMATION:  
APPLICANT: Barnwell, John  
TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,  
TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Darby and Darby  
STREET: 805 Third Ave.  
CITY: New York  
STATE: New York  
COUNTRY: USA

ZIP: 10022-7513  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/719,821C  
FILING DATE: 09/30/96  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Gogorié, Adda

REGISTRATION NUMBER: 29,714  
REFERENCE/DOCKET NUMBER: 5986/17686US3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 527-7700  
TELEFAX: (212) 753-6237

TELEX: 236687

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3337 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium vivax

IMMEDIATE SOURCE:

CLONE: FVMB3.3.1

US-08-719-821C-1

Query Match 0.6%; Score 61; DB 3; Length 3337;  
Best Local Similarity 42.3%; Pred. No. 0.013;  
Matches 456; Conservative 0; Mismatches 615; Indels 6; Gaps 2;

QY 5692 ACTCAAAATGTCAGAAATGCTTTGGTAAAGCAAAATCTCATTCATTAAACAATT 5751  
DB 504 ACAGAGAAGAGACTACCTGGTATATGAGAGAGAGAGAAAGCTGGCGATATGGAAGCA 563  
QY 5752 GGAATTGAGCTGAGATTGAGAGCTGAGAGAGCTGAGATGACAGCTTCTGTAGCAGTGAAT 5811  
DB 564 GGGGAGAAAGCTGGTGAATTGGAGACAGGGAGAAAGCTGGCGATTGGAGACAGGGAA 623

```
QY 5812 AAGATTGTAATATATACATGACAGATTAAATCATGCAAAAATCACTGCGAAGGAAAT 5871
Db 624 GAACCTGGCGATTTTGAAGACAGGGGAAAGAGCTGGTGTATTGGAGACAGGGGAAAGAACT 683
QY 5872 GTCCGAGTTATTAACAGAGCTGATGCGGTAATTTGTATTAATGAGGAAACAGTGTCTGA 5931
Db 684 GCGGATTTGGAAGCAGGGGAAAGAACTGAGATGCGGAAATGAAAGAGCAACTGGA 743
QY 5932 GTGCCCCGTGACAGATAGAGACCTCAACCACTGTGAA--TGAATTAACAGATCTACA 5988
Db 744 GATGCGGAAACTGAAATATGAGCAACTGTGTATGTAGACAGAAAGATAGTTCACTGAT 803
QY 5989 AAAGCATATGTAAAGATTTCTACAGTATGCTTAAAGAAAGAAACAGATATATATTAAT 6048
Db 804 GGAGCAGAAAAAGTCAATGTTCTCTGCTCAAGAAAAATGTACAACCTGCCATATGATGAT 863
QY 6049 ACTCAAGGCGCACTGATTAAGTGTAGTAAAGTATTAAGTAAATCTTAATTTAAACGA 6108
Db 864 GCCCTCTTTGGAAGTATTTTGAATTAAGATTAATTTTGAATCATATTAAGATTTCCAG 923
QY 6109 GACTTATCAAAAAAAGAAAAATTAAGTATTAATAAAGATTTGTTACCAATAGTTCAGCT 6168
Db 924 CCACTATTCGAAACAAATTTGTGCGGCTACTGCTAAACATGTTACGGGACAAAGATTTGCCA 983
QY 6169 ACTCATCTTTTAAATCTTTTATTTGGCAATGCGCGCTGTTCAAGACAAAGCCGAG--TG 6225
Db 984 ATGAACCTGTACATTAACAGTGGCAGAAAGCCGCGCAAGTACACAGCGAAAGATTA 1043
QY 6226 GCAAGAACCTGTATATCAACAGTTTATGAGAAACAGAACTCTTGTAAATAATTTCT 6285
Db 1044 GATGCACTCCAGAGGATGATTCGATTGATGTTACAGAACTCCAGAGAAAGTAGAA 1103
QY 6286 ATATTAAATGCAAAACATTATTTCTGTAATAATCAGAGATTAACAGATTTCAATCCGAGTA 6345
Db 1104 TTAGTTTATGATGAAAGGCACTGAAAGAAATCAACGAGAGTGGACCAACGAAAGAA 1163
QY 6346 GTAGTTCTGTGTGTTGGTGGTGAATGTAGAGTAGAGCTTTCTTGATACCAATATT 6405
Db 1164 GGACCAACCGAAGAAATTAAGATGCACTCCAGAGAGTAGATTTCCGATTAAGCAAACTGC 1223
QY 6406 ATTAATAAGAAATTAACAGACAGAGTTGGAATAAATCAATGCTGATGAAGTTTCCGA 6465
Db 1224 AGAAGAGAAACAGAAAGAACTGAGGAGAAAGAAACAGAAAGCTCCAGAAAGAGAA 1283
QY 6466 GAAAGAGCTGAATTAACAGAGATTTCAAGCAAGAAATTTCTTTTGAAGTCGAGTC 6525
Db 1284 GTATTCAGAAAGAACTCCAGAGAGAAAGAGATTTAGAGCAACTCCAGAGATGATTTTC 1343
QY 6526 GCAGCAGCCGCGGTAGAGCCGAGTGGCAGAAACCGTTTCCGTAATCAATTTGCAGGA 6585
Db 1344 GCATTTAGATGAACTACATTAGAAAGAAACCGAAGAAATGACAGAAAGAAACCGTA 1403
QY 6586 AAGACGGAAGTATGATGTGAAAGAAAGCAAGATTTTGTATAAATAAAGCTGATTAACGA 6645
Db 1404 GAGGGAAGAAAGAAACCGTAGAGGAGAAAGCCGTAGAGGAGAAAGAGCTGCAAGAGA 1463
QY 6646 AAACGTTATAGTTCTGTGCAATTTGAAATGCGCAGAGCTGAGTGGCTGCAAAAGAGCT 6705
Db 1464 GAAAGAAAGTATGAGGCACTCCAGAGATGATCTTCAATTTAGAAAGAACTACAGAGAA 1523
QY 6706 GGAATTTGAGCAGCAGTGGCAGTTACCAAGATGAATCAAAACAGAGAGCAAGAGTG 6762
Db 1524 GGAAGAGGGGAGAGAGAGAGAAAGGAGAAAGAGAAAGCGTTAGTAGACAGTG 1580
```

```
RESULT 25
US-09-754-947-2
; Sequence 2, Application US/09754947
; Patent No. 6828110
; GENERAL INFORMATION:
; APPLICANT: Lee, Bruce Andrew
; APPLICANT: Florea, Becky Mar
```

```
APPLICANT: ValKire, Gunars Edwin
; APPLICANT: Biocore Diagnostics, Inc.
; TITLE OF INVENTION: Assays for Detection of Bacillus Anthracis
; FILE REFERENCE: 014907-003310US
; CURRENT APPLICATION NUMBER: US/09/754,947
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,901
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.1
; LENGTH: 2370
; TYPE: DNA
; ORGANISM: Bacillus anthracis
; FEATURE:
; OTHER INFORMATION: surface array protein (SAP)
US-09-754-947-2

Query Match 0.6%; Score 60.2; DB 3; Length 2370;
Best Local Similarity 41.2%; Pred. No. 0.017; Indels 3; Gaps 1;
Matches 501; Conservative 0; Mismatches 713;

QY 5352 AAAAAATATCAATGTGATGCAAAAGATTATATATGACCAATATCATAGCAGTGGAGT 5411
Db 969 AAAAGTACTTCACTACTGTATTAAGCTGTTATATTAAGAAAGCGTTAAAGTAGTGTGA 1028
QY 5412 TGGAGAGCAAAAGAGCCTCTGTGCAAGAGCTTCTGCAAGTACTTACTTGAATTAAGAC 5471
Db 1029 AAGTAAAGAAAGTAAAGTTTCTGCGAAGGTGCTGCAAGTACTTCAATCTTCACTGAGC 1088
QY 5472 AGTTCTTCTCATATGTGATCAAACTGATATTGACAAAGATTGAGAGAAATTAATG 5531
Db 1089 AGTTCAAGAAACAAATTAAGCTGACTTACTTCTTAAGATTTCAACAAACAAATTAAGT 1148
QY 5532 AAATAAGAAAGCAAAATGTTAATGTTCTGACGTAATAATGAGTCAAGTGTGACAAA 5591
Db 1149 ---TTACAGAGCGCAACAGCTTACGTTCAAGTAAATGAAAGATTCATTTAAGCAGT 1205
QY 5592 TCGCAGCTGTTTCCGAGCAAGTGAACAGCTGCAAGTGAAGCTGAGAGTGAAGTAA 5651
Db 1206 AACCACTGAAAGTGAATATGAGTCTGTTAAACACAGAAAGTGTGTTAGTGAATTAAGC 1265
QY 5652 TAAATTAACAAATATCTTTCGACATTAATAAATAATGACTCAAAATGTACGAATGC 5711
Db 1266 TACTGTAAAGTAACTGTATTAATTCGACAGAAAGCAACAGTAAAGTAACTGTAAAGAA 1325
QY 5712 TTTGGTAAAGCAAAATCTCATTCATCTATTAATAAATAATGGAATTTGAGCTGAGTTG 5771
Db 1326 TTCAAAAGTAAAGCACTGTTTTCACACAGTGAATTTGAAGCTTTGCTCAAAAGC 1385
QY 5772 AGCTGAGAGCTGAGTGAACAGTGTCTGTAGCAGTGAATTAAGTGTAAATTAATTCAT 5831
Db 1386 AATGAAGACATTAATTTGAATAAATACTAACCTAGCGCTTTCTACAAAGATGTACAGA 1445
QY 5832 AGCAGAAATTAATCAATGCAAAATGACGCGAAGGAAATGTCCGAGTATTAACAGATC 5891
Db 1446 TTTAAAGTAAAGCTCCAGCTGATCAATACGTTAAAGGTTTACAGCTCTGTAAAC 1505
QY 5892 TGATGCGTAAATTTGCTAATTTATGAGAAACAGTGTCTGAGTGGCCCGTGCAGCAATAG 5951
Db 1506 AGTGAAGTAACTGTGATTAAGATGTAAAGAAATTAAGAAACAAATAATTTGAAGCTTAA 1565
QY 5952 AGCTCAACAGTGTGAATGAATTAACAGATCTACAAAGCATATGTAAAGATTTCTAC 6011
Db 1566 TGTGAACAGAGATTAATGTTCTGAATGACAGAGTCAAGAAAGCTGTGAATTAATACAGT 1625
QY 6012 AGTGAATGCTAAAGAAAGAAACAGATGATTAATTAATCAAGGCGCAAGTGAATAGT 6071
Db 1626 ATTAAGTGAATAATCTGTGTAAGAAAGAAAGCAAAAGCTAATTAAGCTTGTAAATTAAG 1685
QY 6072 GGTAGATAAGTATTAATAAATCTTAATTAATTAACGAAAGCTTAATCAAAAGAAAGAA 6131
Db 1686 TCCAGGTGATCTCTTAATTTGAAGTGTGTTTGAACAGAAATTAATTAATTAATGAT 1745
```

QY 6132 AAGTATATAAAGGATTGTTTACCAATAGTTCAGCTACTCATCTTAAATCTTTATT 6191  
| | | | |  
DB 1746 TACTAGAGAAAACAAAAGAAATGCAATGCTGTTTCACTTCTCTGTAAGATGCAATAG 1805  
| | | | |  
QY 6192 GCGAAATCCGCTGTTTCAAGCAAGCCGGAAGTGCAGAACTGTTAAATATCACAAGT 6251  
| | | | |  
DB 1806 ATTGATATTTAAAGGTGCAAGAGCTGAACTAAAGATTAACACACAAACAAAGAAAG 1865  
| | | | |  
QY 6252 TTATGAGAAACAGAGCTCTTGTAGAAATTTCTATTTAAATGCAAAACATTTATCTGT 6311  
| | | | |  
DB 1866 TAAAGAGTAGCAGCAATGATGCAAGATTAAGTTAACTGTAAGTAAATTAATTAATCTGT 1925  
| | | | |  
QY 6312 AAAATCAGAGATTACAGCAATTCATCGAGTAGTAGTTCTGTGTTGTGTGAGAA 6371  
| | | | |  
DB 1926 TGTGCAAGGTGCAAAAGCTGAGTGAAGTATTAAGTTAAAGTTGATGTAAGTAAATTT 1985  
| | | | |  
QY 6372 TGTAGGAGTAGAGCTTTCTTGTGATACCAATTTATTTAAAGAAATACCAAGACAGAGT 6431  
| | | | |  
DB 1986 AATCAACAATCATTCATTCAAAGTTGTGATACAGACCAACTGCTAAAGATTAGCAGT 2045  
| | | | |  
QY 6432 TGGAAATCTAATGCTGATGAAGGTTTCGAGAGAGAGCTGAATTAACAGCAGATTC 6491  
| | | | |  
DB 2046 AGAATTTCAAGCAATCTCTTAAAGAAATGCTCCAAATGCTGATTTAAAGCTGCACT 2105  
| | | | |  
QY 6492 TAAAGCAAGAAATTTCTCTTTTGAAGTCGAGAGTCCAGCAGCCGAGGTAGAGCCGAGT 6551  
| | | | |  
DB 2106 TTTAAATCTTATCTGTTGATGTTGATGCTGCTGCACTACAGAAAGCAAGCTTTCA 2165  
| | | | |  
QY 6552 GCGAGAAACCGTTTCCG 6568  
| | | | |  
DB 2166 TGTAGAAATTTGTTCTG 2182  
| | | | |

## RESULT 26

US-08-559-896B-1  
; Sequence 1, Application US/08559896B  
; Patent No. 6310046  
; GENERAL INFORMATION:  
; APPLICANT: Patrick E. Duffy  
; APPLICANT: Christian F. Ockenhouse  
; TITLE OF INVENTION: SEQUESTRAIN  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John Moran  
; STREET: USA MRC - MCMR-JA  
; CITY: FORT DETRICK, FREDERICK  
; STATE: MARYLAND  
; COUNTRY: USA  
; ZIP: 21702-5012  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh 7.5  
; SOFTWARE: Microsoft Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/559,896B  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Moran, John  
; REGISTRATION NUMBER: 26,313  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 619-2065  
; TELEFAX: (301) 619-7714  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1956 base pairs  
; TYPE: Nucleic acid

STRANDEDNESS: Double  
TOPOLOGY: Linear  
US-08-559-896B-1

Query Match 0.64; Score 60; DB 3; Length 1956;  
Best Local Similarity 45.18; Pred. No. 0.017;  
Matches 222; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

QY 459 AGATGATTTTAAAGAGCTTTGGAAGAGCCAAACATGTTAAAGTTTAAATGGAATCAT 518  
| | | | |  
DB 1392 AGACTATTTTACAAAGAAATTTAAAGACAAATGTTAAATTTTCCATATAGTGA 1451  
| | | | |  
QY 519 TCCAGTAGATGAAAGTAAAGTAAATTCATGGAATCCGATGGAAGATTACGATGAAG 578  
| | | | |  
DB 1452 TTTTAAATTCAAAGATGAGAAAGATTAACACCAATTCATCATTAAGAGAAATCA 1511  
| | | | |  
QY 579 AAAATCAATGCTTTGGAAGCATCGTTTATATGCGCGGATTTAGATTGAAAGATAC 638  
| | | | |  
DB 1512 AGTAGATGTTGTGCGAAGAAATATTCAGATTATTCAGAGGATATATTAAGG 1571  
| | | | |  
QY 639 TGCATATCTAAAGACAGGAATTAAGATTAAATTTAGTCAATTTAGTGAATGCAAT 698  
| | | | |  
DB 1572 CCAAGAGATTAACCTGAATATGTTAGATATTAAGAAATTAACAAATATGATATTA 1631  
| | | | |  
QY 699 AAATTCGCTGACCGAGATTTAAAGCTAACCAAGCAAAATCTGAGATTAATTTCT 758  
| | | | |  
DB 1632 AAATGTTGATGATATTAAGAAATGTTGAGATATTAAGGTTGAGATATTAAGG 1691  
| | | | |  
QY 759 TTCAGCTCACATGATTTCTCTCAAAAAGCTATGAGAAAAATTCACACTGTTGAGAAAG 818  
| | | | |  
DB 1692 TGATGATATTAACAATGTTGATGATTAAGAAATGTTGATGATTAAGAAATGTTGATG 1751  
| | | | |  
QY 819 AATGAGAAATATGTAAGAAAGAAATTAACCAAGCAATATTAATCTGATGCTGATTTGA 878  
| | | | |  
DB 1752 TATTAAGAAATGTTATGATATTAACAATGTTGAGATATTAACAATGCTGAGATTA 1811  
| | | | |  
QY 879 AGCAGATGAAATATTAAGAAATTTAGTGCAGAACTACAAATGGAAGATTTATTAAGAA 938  
| | | | |  
DB 1812 TAATGCTGAGATATTAACAATGTTGAGGATATTAACAATCTGTATATATCAAGCT 1871  
| | | | |  
QY 939 AGGGGAAAAAGA 950  
| | | | |  
DB 1872 TGAACATATAGA 1883  
| | | | |

## RESULT 27

US-09-351-794A-1  
; Sequence 1, Application US/09351794A  
; Patent No. 6641815  
; GENERAL INFORMATION:  
; APPLICANT: DUFFY, PATRICK E.  
; APPLICANT: OCKENHOUSE, CHRISTIAN F.  
; TITLE OF INVENTION: SEQUESTRAIN  
; FILE REFERENCE: 38644-1/7519  
; CURRENT APPLICATION NUMBER: US/09/351,794A  
; CURRENT FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 08/559,896  
; PRIOR FILING DATE: 1995-11-17  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1956  
; TYPE: DNA  
; ORGANISM: Plasmodium falciparum  
US-09-351-794A-1

Query Match 0.64; Score 60; DB 3; Length 1956;  
Best Local Similarity 45.18; Pred. No. 0.017;  
Matches 222; Conservative 0; Mismatches 270; Indels 0; Gaps 0;

QY 459 AGATGATTTTAAAGAGCTTTGGAAGAGCCAAACATGTTAAAGTTTAAATGGAATCAT 518  
| | | | |  
DB 1392 AGACTATTTTAAAGAAATTTAAAGACAAATGTTAAATTTTCCATATAGTGA 1451  
| | | | |

QY 519 TCCAGTAGAGAAAAGTAAATTCATGTAATCCGAATGGAAGCAATTACGCTAGAAAG 578  
DB 1452 TTTTAAATTCAGAAAGTAAAGAAATTAACACACCAATTCAGCAATTAAGAAAGAAATCA 1511  
QY 579 AAAAATCATGCTGTGAAGCATCGGTTTATATGCGCGGATATTAGATTGAAGATAC 638  
DB 1512 AGTAAGTGTTCAGAGAAAATAATTCAGATTATTCAAGAGATTAATTAAAAAATTAAG 1571  
QY 639 TGCATACTAAAGACAGAAATTAAGATTTTAAAAATTTAGCAATATTAGTCAAT 698  
DB 1572 CAAAAGATACACCTGAAATGTTAGATTAATATAGAAATTAACAAATATTGATATAA 1631  
QY 699 AAATTCGTCTGACCGGAGATTTAAAGCTCAAGACAAATCGAGATTTATTCT 758  
DB 1632 AAATGTGTGATATATAAAAAATGTGGAATATTAAGTGTGAGATATATAAAGGT 1691  
QY 759 TTCAGCTCATAGATTCCTCAAAAAGCTATGGAAAAATTCAGCTGTGAAAGAG 818  
DB 1692 TGATGATATAACAATGTTGATGCTATTAATAAATGTTGATTAATAAATGTTGATG 1751  
QY 819 AATAGAGATATGTAAAGAAATACCAAGCAATATTGAATGCTGATGTTGA 878  
DB 1752 TTTAAAAATGTTGATGATTAACAATGCGAGATATAACAATGCTGAGATACAA 1811  
QY 879 AGCAGATGAAATATTAATAATTAGTCCGAAGCTACAATGGAGATTTATTAAGAA 938  
DB 1812 TAATGCTGAGATATTAACAATGTGGCGATATTAACAATCTGTAGATATATACAACT 1871  
QY 939 AGGGGAAAAAGA 950  
DB 1872 TGAACATATAGA 1883

RESULT 28  
US-09-806-708B-22  
Sequence 22, Application US/09806708B  
Patent No. 6784342  
GENERAL INFORMATION:  
APPLICANT: The University of British Columbia  
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants  
FILE REFERENCE: 4810-58741  
CURRENT APPLICATION NUMBER: US/09/806,708B  
PRIOR FILING DATE: 2001-04-03  
PRIOR APPLICATION NUMBER: US 60/147,133  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 22  
LENGTH: 1141  
TYPE: DNA  
ORGANISM: Artificial sequence  
FEATURE:  
NAME/KEY: promoter  
LOCATION: (1)..(1141)  
OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. PAEI promoters  
US-09-806-708B-22

Query Match 0.6%; Score 59.4; DB 3; Length 1141;  
Best Local Similarity 10.4%; Pred. No. 0.019;  
Matches 80; Conservative 287; Mismatches 405; Indels 0; Gaps 0;  
QY 858 TGAATCTGATCTGTATTTGAGACAGATGGAATATATAAATTAAGTCGAAAGTACAA 917  
DB 94 TSBRYABARKDKMTIVBMTMTKMGKTGRNHYMRABVDVDDHNYLVAANNAWTTC 153  
QY 918 TGGGAGATTATATAAGAGAGAGGAAAAAAGAACTTATACATCTCTTAAGTTATC 977  
DB 154 MMKDXDKRTMMWKNNATGMDDDTKYHMMNNNGBVTYVMVRYUKTRDMSBRNNYGC 213  
QY 978 AGATGTGAAGCTTCGCTAGAGTAATAAGAAAGATGATGAAAGAAATGTCACAT 1037  
DB 214 BMMKMSYDVTTYTWVDDMKCKKVRWVTRGRMRWYVAMBTARRRNNNGTBMAY 273

QY 1038 TACAGCTGAAGCAAGAATTTCTATGATGCAACTTAGTACTAGCTTGCAAAAGCACTC 1097  
DB 274 RRTMTNNNNNNNAKMKCAKATGWRBRYNSTCTTMSKITKATSCANNCRADNDKD 333  
QY 1098 TTTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAATGATTTTAAAGTTATT 1157  
DB 334 HKMMKMSAAMGVNN 393  
QY 1158 GACAAGTATGCTCCAGTGTCTGTTATTTGAAAAAGATGCCAAATCGAAGCAACGAAGAA 1217  
DB 394 RKTWNNNNNGTMMKRWMAWMTKMDMBGTYNNNNNGRTTYGATTKKKMTYYKWTAN 453  
QY 1218 GGCAATATTCATCTTACAGTGAAGTGAAGCAACTGAGAGACGATCTTCATCATT 1277  
DB 454 NCKRMAWDHCTCHNNTTMMKTKYNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 513  
QY 1278 AAAAATTACCAATTTATATTTGAGAAAGCCAAATGAAACTTCTCAGTATCGAGCGG 1337  
DB 514 NNNMMDYMMKACTMYKYBVCCKMNNVYAAWYTKSMYTSRYRWKTNNSWRBSDTBSM 573  
QY 1338 ATATATTTCTGCAAAAAGTAATTCATATGTAATTTGAAGAGAAAGTAAATGCAAGG 1397  
DB 574 GRANNYABABHYGYKMTNRMWBSHTWBHBRAGAAYMBWMBYBAKCHCMRAWYKAKYA 633  
QY 1398 AAGAGCAGATATTAATCTCAAAATCGAAATGTAATGATGCTGCTGTTGGAAC 1457  
DB 634 GAGGNN 693  
QY 1458 GATGAGAGATTCATATAAGTAGCTCTTCAATGTTGTCGAGAGAGAAATTAATC 1517  
DB 694 MWGCMNNAITDTRITMMKNN 753  
QY 1518 TTCGCTCAAGATGCTTAAGAGCAAAAAGTGAATCAGAAACGATGATGTAATGTGAG 1577  
DB 754 KWAMGMAADAAABTTDKRNGAYTKYTTNNNNNTYRGVNTTARDDMANNNNNNNNNN 813  
QY 1578 AAGTAAACGATTAATTCATTCGAGCTGCTGTAAGAAGTGGATGGGGAT 1629  
DB 814 NNNNGMSDMWVWYVANYGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 865

RESULT 29  
US-08-682-517-7  
Sequence 7, Application US/08682517  
Patent No. 5874267  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Expression of surface layer proteins  
NUMBER OF SEQUENCES: 25  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/682,517  
FILING DATE:  
CLASSIFICATION:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4197 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

Query Match 0.6%; Score 59.4; DB 2; Length 4197;  
Best Local Similarity 42.0%; Pred. No. 0.03;  
Matches 826; Conservative 0; Mismatches 1111; Indels 30; Gaps 7;  
QY 5830 ATAGCAGATTAATAATCATGCAAAATATCACTGGAAGGAAATGTCGAGATTATACAGAG 5889

Db 1916 AAAGAAAACCAATCTGTGTAAACCAATGATTAAGGAAAGCTACATTGGTATCCGT 1975  
 Qy 5890 TCTGATGCGGTATTTGCTATTTATGAGGAAACAGTGTCTGAGTGCCCGTGCAGCAATA 5949  
 Db 1976 TCTGATGAGTAAAGATTAATGCAACCAATTTGCTGGATTGATTTATATCTTCTGAT 2035  
 Qy 5950 GGAGCCTCAACCAAGTGTGAATGAAATTCAGAGATCTCAAAAAGCATATGATTAAGATCT 6009  
 Db 2036 GCAAAAACAAAGCGAGCTTGATGAAGTGAACCAAAAGCAGTTGACCAATCTTACTTC 2095  
 Qy 6010 AAGAGATTGCTAAAGAAAGAAACAGATGATTAATTTACTACTCAAGGCAAGTAAATAA 6069  
 Db 2096 CAAGACCATATCTTGATGGCTCAGCTATCAAGCATACAAAATTCAGATCTTATATAA 2155  
 Qy 6070 GTGAGATTAAGATTTCAAAAATCTTATATTAAGAAAGCTTATCAAAAAGAAAA 6129  
 Db 2156 GCTGTACTAGTTGATGGTCTGAAACTGCAGATTTTGGCAGCAAAATTAAGTAAACCA 2215  
 Qy 6130 ATAAATTAATTAAGAAAGTTGTTCACATAGTTGAGTCTCATATCTTAAATCTTTA 6189  
 Db 2216 AGCGCAAAAAGTAACTGTACTTCTATTAAGAAAGCACTTATACATCTACAAATCT 2275  
 Qy 6190 TTGGCAATGCGCGTGTGTTCAAGACAGCGGAGTGGCAGAGCTGTTAATTCACAG 6249  
 Db 2276 GGTGTAAATGATTAATTAAGTATACCAAGTTATCTCCAAATCTGAGTACAGATA 2335  
 Qy 6250 GTTTATGAGAAACAGAAAGCTCTGTGAGAAATTCATATTAATGCAAAAACATTAATCT 6309  
 Db 2336 ACTTATGAAGCTACTTTATCTTCTACAGAACTGTTATTAACCTGCTAAGATTTAGAA 2395  
 Qy 6310 GTAAATCAGAGATTAACAGAAATTCATCGAGTAGTAGTTCTGTTGGTGGTGA 6369  
 Db 2396 GTTACTCTCAGTGAGTGTAAACCACTGCTGTTAAAGTAAATGTTAGTACAGGTATGCTGT 2455  
 Qy 6370 AATGTAGAGTAGAGCTTCTCTGTATACCATATTTATTAAGAAATTCAGAGCAAGA 6429  
 Db 2456 AATACAGACGTAAAGATATGATCTTATCTGCTAAAGAGCTACAGCTACATTCACACT 2515  
 Qy 6430 GTTGGAAAACTACATGCTGATGAAGGTTTCGAGAAAGAGCTGAATTCACAGAGT 6489  
 Db 2516 ACAATGAAAGTTCCAACTCTTACACTGTGTGACTACAA--TTCAATACACTGAT 2572  
 Qy 6490 TCTAAGCAAGAAATTTCTCTTTTGGAGTCGAGTCGAGCAGCCGGGTAGAGCCGGA 6549  
 Db 2573 TCTGGTTCAACAGCACTCTATTTGGTGTGCTGTAAGAAACCCAGTGAATATGCTGT 2632  
 Qy 6550 GTGGCAGAAACCGTTTCCGTAATCAATTTGCAAGAAAGACGAAATGATGTGAAGAA 6609  
 Db 2633 GTATCAGGCAAAACATTAATTAATCTCGAGCTAAATGTAATGAAGTATTTGGTAAACG 2692  
 Qy 6610 GCA---AAGATTTTGGTAAAAAAAGCTGAGATTCAGCAAAAAGCTTAATGTTCTGTTGC 6665  
 Db 2693 GATGAGGAAGCATTTATTAATCAATATGCACTGAAGGCCAAAAGTAAACATTCATAT 2752  
 Qy 6666 AA--TTGAAATGCGCAGCTCGAGTGCTGCAAAAAGAGCTGAATTTGAGCAGCAGTG 6723  
 Db 2753 AATGTAGATGTGATACAGTTACATTTAAGTAATTAAGTGTGTTAATTTCTTCAACTGA 2812  
 Qy 6724 GCAAGTTACCAAGATGATCAACACAGAGCAAGATGAATAATTC-----TAAATTA 6778  
 Db 2813 GCTATCAAAACAGTTGCTCCAAACACAGCAGCTCCCAATCTGCGCATTAACATTA 2872  
 Qy 6779 TGACTGAAACAAAGT-----TAGATGTAATAGCAAAAATGAGATTAATCAAGTACT 6831  
 Db 2873 ACAACGACGCTGTGTGTGTTAGTTGATTTAACAATCTGCACTAACCTTAGAAATTTCA 2932  
 Qy 6832 GGAATCGGTTCAAGCCGAGCTGGAATCTTGGAGCCGAGATCTGAGATGTTCTGTC 6891  
 Db 2933 TTAGCGTATGAGATCTTAATTAATTAAGTGAACAACGTTGATGCAACTGTTTCATTA 2992  
 Qy 6892 AATTAATTTGCAATTAAGTGAAGAAACAGATATGATCATAGTACTTTCATCTTCTACT 6951

Db 2993 AAGATAGTGCAAAATTAATTCATTAATCTCTTAACATTAAGTGAAGAACTGCTAATACAGCT 3052  
 Qy 6952 GATGTAATGTAAGAAAGCTCTTAATTAATTAATTCGAATTCCTTGACAGCCGCTGAGAGCC 7011  
 Db 3053 GTATTTGCTAACAACGTTTCAAGCTGTGATACATTAATCTTCTTAATCTGCTGTGATCACTTAACA 3112  
 Qy 7012 GCAGGCTTGGCAGAGTATCCGAGAGTGTCTTGTTAACACTTAATTAATAGTCTGTGATA 7071  
 Db 3113 GTTACTTAATGCAAGATGCTTAATAAATGCTGACAGTGTGCTGAAAATTAATTAATTAATTAATTAAT 3172  
 Qy 7072 GCTGAGTTCACATTAATCTGATTTGATCTCCGTAAGAGAAATGTAATTAATTAATTAATTAATTAAT 7131  
 Db 3173 ACATTTAAAGAAACCTACTGAGCAATTAATCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3232  
 Qy 7132 AAAGAGAAAAAATTAATTAAGCAAAACAGCAAGCAATGCAAGAAATCGAGAGAGCAATC 7191  
 Db 3233 TCAGCAGCTACAGAGCTGAATATATCTTC-----TAAATCAATTTGCTGCAAGATTAATTAACA 3286  
 Qy 7192 GGAGCCAAATGCTTGTGTAATTAATTTTGGAAACAGCTGTAGAAATGAAAAAATTTCTGAA 7251  
 Db 3287 TTGCAACAGGTGAAGATTCACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3346  
 Qy 7252 GGAAGAGCAAGAGTTTTAAACCTTTAGACGAAGTTAACAAAGAACAAAGATTAATTAATTAATTAATTAAT 7311  
 Db 3347 TTACAGGTAAAGAAAGGTGCAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3406  
 Qy 7312 GTAATGATGCTACGAAAAAAATCTTCAATCAGCAGGATTTTCTACAGAAATCTTCT 7371  
 Db 3407 ACTGCACTGTTTCTGAGAGCAAAAGTATTAATTAATCAGTACAAACAGGTGTGTTCT 3466  
 Qy 7372 GTAAAGCGATGAGAGATTAATCTCAGAGAGAGAAATTAAGCAATTTGGAAGCTTCT 7431  
 Db 3467 GAAATGAGTTACATCTCTTCTGTTAATCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3526  
 Qy 7432 GATATTAATGCAAAAAATGATGATTAATTAACAAAGAGCAAGAAATTAATTAATTAATTAATTAATTAAT 7491  
 Db 3527 GATCAAGCTGTTGCAAGAAACAGCTGTCTCAAAACATTCAGATTAATTAATTAATTAATTAATTAATTAAT 3586  
 Qy 7492 GGTGTTTGGGAATGCAAGCTGCTTCTGCTCCAGATCAGAAACAGTGTGACAAATTAAT 7551  
 Db 3587 GTGGGTGAAAAAGTATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3646  
 Qy 7552 AAAAGAAATTCGAGATTAATCTGTTGAATTTCTTTTGTGAAGAGCTGAAAAAGTAAAT 7611  
 Db 3647 CCAACAGCAATTAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3706  
 Qy 7612 GTTAGATCGATTAATTAACAGAAATGTTGCTTTAACAAGATTAATCAAGGCTGCTGAGAGCA 7671  
 Db 3707 CAAGCTGCAAAATCTGCTGCTAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3766  
 Qy 7672 TTGGAAATGAGAGCTGCTATGCAAGATTAATTAATCTAATGAGAGATCAAAATATCATGATTT 7731  
 Db 3767 TCTGCAACAGGTGCTTACTATTAAC--ATTAACTTCTACTGTGAACAACAGTATGATTAACA 3823  
 Qy 7732 AAAAATCTTAAGCTATTAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7778  
 Db 3824 ATTACTGAACAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3870

RESULT 30  
 US-08-682-517-8  
 ; Sequence 8, Application US/08682517  
 ; Patent No. 5874267  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: Expression of surface layer proteins  
 ; NUMBER OF SEQUENCES: 25  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
 ; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/682,517  
FILING DATE:  
CLASSIFICATION:  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4197 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Bacillus sphaericus  
INDIVIDUAL ISOLATE: P-1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 95..3850  
FEATURE: mat\_peptide  
NAME/KEY: 185..3850  
LOCATION: 185..3850  
FEATURE: sig\_peptide  
NAME/KEY: 95..184  
LOCATION: 95..184  
US-08-682-517-8  
Query Match 0.6%; Score 59.4; DB 2; Length 4197;  
Best Local Similarity 42.0%; Pred. No. 0.03;  
Matches 826; Conservative 0; Mismatches 1111; Indels 30; Gaps 7;  
QY 5830 ATAGAGATTAAATCATGCAAAATTCCTGGAAGGGAATGTGGAGTTATTAACAGAG 5889  
DB 1916 AAAGCAAAACAATCTCTGTAAACAAATGATTAAGGAGCTACATTTTTCGT 1975  
QY 5890 TCTGATGCGGTATTTGCTAATTATGAGGAACAGTGTCTGAGTGCGCCGTGACAAAT 5949  
DB 1976 TCTGATGAGTAAAGATTAATGCAACCAATTTGCTGATTAATTAATTAATCTGTAT 2035  
QY 5950 GGAGCCTCAACAGTGTGATTAATTAACAGATTTCAAAAGCATGTAAAGATTCT 6009  
DB 2036 GCAAAACAAGCGACTTGATGAGTGAACCAAAAGCAGTGGACCAATCTTACTTC 2095  
QY 6010 ACAGGATTTGCTAAAGAAAGACAGTATTAATTAATTAATTAATTAATTAATTAAT 6069  
DB 2096 CAAGCAACATATCTTGATGCTCAGCTATCAAAAGCATTAACAAATTCATTAATA 2155  
QY 6070 GTGTAGATTAAGTATTCAAAATCTTAATATTAAAGAACTTAACAAAAGAAAA 6129  
DB 2156 GCTGTACTAATGTTGATGTTCTGAACTGCAAGTATTTGAGAGAAATTAATTAACAA 2215  
QY 6130 ATTAAGTAAATAAAGATTTGTTACCAATAGTTACGTAATCTTAATTAATTAATCTT 6189  
DB 2216 ACCGCAAAAGTAATCTGTACTTCTTAAGAAAGCAATTAATCAATCTACATACT 2275  
QY 6190 TTGGCAAAATCCGCTGTTCAAGACAAAGCGGAGTGGCAAGCTTAATTAATCAAG 6249  
DB 2276 GGCTTAATGATTAATTAAGTAAAGTAAACCAAGTTATCTCAAAATGTAGCTACAG 2335  
QY 6250 GTTTATGAGAAAAACAGAGCTCTTGTAGAAAAATCTATATTAATGCAAAACATTAAT 6309  
DB 2336 ACTTAAGAGCTACTTATCTTCTACAGAACTGTTATTAACCTGCTAAGAAATTAAGA 2395  
QY 6310 GTTAAATCAAGAGATTACAGAAATTCAGAGTATGATGTTCTGTTGGTGTGGA 6369  
DB 2396 GTTACTTCAGTGTGATTAACCACTGCTTAAAGTAAATGCTACAGGTAATGCTGTT 2455  
QY 6370 AATGTGAGATGAGAGCTTCTCTGATTAACCAATTAATTAATTAATTAATTAATTAAT 6429  
DB 2456 AATACAGACGTAAGATTAATGATTAATGCTTAAGAAAGCTACACTACATTCACAGCT 2515  
QY 6430 GTTGAAGAAATCAATGCTGATGAAGTTTCGAGAAAGAGCTGAATTAACAGAGAT 6489  
DB 2516 ACAAATGAAGTTCCAAATCTTACACGCTGTAGCTACAA---TTCAATACAGCTGAT 2572  
QY 6490 TCTAAGCAAGAAATTTCTCTTTTGAAGTCGAGTGCAGACCGGGGATGAGACCGGA 6549

DB 2573 TCTGTTCAAAACGCACTATTTGTTTCTGTAAACCAACGAGGAATATGCTGCT 2632  
QY 6550 GTGCGAGAAACCGTTTCCGTAAATCAATTTGCAAGAAAGACGAGTATGATGGAAGAA 6609  
DB 2633 GTATCAGGCAAAACATTAATTAATCTTGTGAGCTAATGATTAATGAATTTGGAGACG 2692  
QY 6610 GCA---AAGTTTGTGTAATAAAGCTGAGATTACAGCAAAAGGTAATATGTTCTGTTGC 6665  
DB 2693 GCATGGGAAGCATTAATTAATCAATATGCACTAAGGCGCAAAAGTAAATCATATAT 2752  
QY 6666 AA---TTGGAATGCGCGAGCTGAGCTGCAAAAGAGCTGAATTTGAGACAGAGTGC 6723  
DB 2753 AATGTAGTGTGATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2812  
QY 6724 GCAGTTACCAAGATTAATCAACAGAGCAAGAGTGAATAATTC---TAAATTA 6778  
DB 2813 GCTATCAAAACAGTTGCTCAACCAACAGAGCTCCAACTAGTGGCGCATTAACATTA 2872  
QY 6779 TGACTCGAAACAGT-----TAGATCTAATACCAAAATGATTAATTAATCAAGTACT 6831  
DB 2873 ACACCAAGCAGCTGTGTGTTAGTTGATTTAACTCACTGCACTAATTAATTAATTAAT 2932  
QY 6832 GGAATCGGTTGAGCGGAGCTGGAATTTCTGAGCGGAGTATCTGAGTGTCTGTC 6891  
DB 2933 TTAGCTGTGAGATCTTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 2992  
QY 6892 AATTAATTTGCAAAATTAAGTAAAGCAATATGATTAATGATTAATGATTAATGATTAAT 6951  
DB 2993 AAAGTATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3052  
QY 6952 GATGTAATGTAAGATCTTAATTAATTTTCAATTTCTTGAACCGCGTGGAGAGCC 7011  
DB 3053 GTATTGTCAACCTGTTCAAGCTGTGATTAATTTCTTAACTGCTGTGATTAATTAAT 3112  
QY 7012 GCAAGTCTTGAGAGAGTAAACGAGTGTGTTCTGTTAACAATTAATTAATTAATTAATTA 7071  
DB 3113 GTTACTTAATGAGATGCTTAATAATGCTGAGTGTGCTGAAATTAATTAATGCTGATTA 3172  
QY 7072 GCTCGAGTTCAAAATTAATCTGATTAATTTGATTTCTGTTAACAATTAATTAATTAAT 7131  
DB 3173 ACATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3232  
QY 7132 AAAGGAAATAAATTAATTAAGCAACAGAGCAATGCAAGATTTGAGAGAGCAGCAATC 7191  
DB 3233 TCAGCAGCTACAGCAGTGAATTAATCTTC-----TAAATCAATGCTGAGATTAATCA 3286  
QY 7192 GGAGCAATGCTGTGTAATTAATTTTGAACAGCTGTAGAAAGATTAATTAATTAATTAAT 7251  
DB 3287 TTTGCAAGAGTGAAGATTTCACTTTTAATTAATTAATTAATTAATTAATTAATTAATTA 3346  
QY 7252 GGAAGGAAACAGAGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7311  
DB 3347 TTAGCAGGTAATAAAGTGTGCAACAGGTGATGATTAATTAATTAATTAATTAATTAAT 3406  
QY 7312 GTAAATGATGCTACGAATAAATTTTAATTAATTAATTAATTAATTAATTAATTAATTA 7371  
DB 3407 ACTGCACTGTTTCTGAGAGCAAAAGTATTAATTAATTAATTAATTAATTAATTAATTA 3466  
QY 7372 GTPAAAGGGAATAGAGATTAATCTCAGGAGAAAGAAATTAAGCAATTTGTAAGACTTCT 7431  
DB 3467 GAAAGTGAAGTTACATTTCTCTCTGTTAATTAATTAATTAATTAATTAATTAATTAAT 3526  
QY 7432 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7491  
DB 3527 GATCAAGTGTGAGAGAAACAGCTCTCAAAAGCATTAACAGATTAATTAATTAATTAAT 3586  
QY 7492 GGTGTTTGGCACTGCAAGTCTTCTGCTGATTAAGAAACAGTGTGCAATTAATTAAT 7551  
DB 3587 GTGGGTGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3646  
QY 7552 AAAGAAATTCGAGATTAATCTTTGGAATAATCTTTTGTGAAGAGCTGTAATAATTAAT 7611

Db 3647 CCACAGCAAAATACATTCGTAGTGAATCTGCTGTAATACATTAGCTTCAGTAC 3706  
Qy 7612 GTTAGATCGATATATACAGAAATGTGCTTTAAACAGATATCAAGTCTGTAGAGCA 7671  
Db 3707 CAAGCTGCAAAATCTTGTCTGTCAATTTACTTTAAACACTGCAGATTAAGTTACACT 3766  
Qy 7672 TTGGGAATGAGAGTGCCTTAGCAGAAATTAATTTCTAATGAGATCAAAATATCAGTAT 7731  
Db 3767 TCTGCAACAGAGGTGCTACTATTAC---ATTAACTTCTAGTAAACACAGTAGTACTACA 3823  
Qy 7732 AAAAATTTCAAGCTATAGCAAAAATATGATGTATTATGTAAGA 7778  
Db 3824 ATTACTGAACCAAGTATTAATTAATAATGAATTAAGCAATTAAGTAAATA 3870

## RESULT 31

US-09-601-198-56/c  
Sequence 56, Application US/09601198  
Patent No. 6531583  
GENERAL INFORMATION:  
APPLICANT: Caseall, Gail H.  
APPLICANT: Chen, Bileon Y.  
APPLICANT: Glass, Jennifer S.  
APPLICANT: Glass, John I.  
APPLICANT: Heiner, Cheryl R.  
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA  
FILE REFERENCE: UAB-13452/22  
CURRENT APPLICATION NUMBER: US/09/601,198  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: 60/073,189  
PRIOR FILING DATE: 1998-01-30  
NUMBER OF SEQ ID NOS: 181  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 56  
LENGTH: 14066  
TYPE: DNA  
ORGANISM: Ureaplasma urealyticum  
US-09-601-198-56

Query Match 0.6%; Score 59.4; DB 3; Length 14066;

Best Local Similarity 41.7%; Pred. No. 0.044;

Matches 436; Conservative 0; Mismatches 606; Indels 3; Gaps 1;

Qy 125 CGACTGAGATTTTGGACAAAATAGAAAAAGGATTAAGTATGACATTAAGTACTACAA 184  
Db 9146 CTAAATATAGCTATTTAAACAGTTAGTGCACATTAAGTCAATTTAAAGCTTAATA 9087  
Qy 185 ACAAGATTCAGGGAGAGACGCTTTTAAACAGTTTAAATAGATTTGCTTTAAACAGAAATA 244  
Db 9086 CAAATGATGAGACATTAAGAAAAGATCAAAATTTGGAAGAGATTTTGGTCCAACTAATA 9027  
Qy 245 ATATGCAAAATCTAATTTTGGGAAAAAGATAGAGGGGGTAAATATCTTTTAACT 304  
Db 9026 ATTTAATGATCAAAAAGATTTGAAGCAAAATTAATTAATCTAATTAATCAATTTAATG 8967  
Qy 305 TTGTCAATGAAAAATTTGAAGTATGAGATTAATCAAGGAATTTGAGAAAAATAATG 364  
Db 8966 AAGGTGAATTAATTAATTTTAAAGGGTTTAAAGAAAGAAAGCAATATCGTTTAACTA 8907  
Qy 365 GAGGAAATTTATTTCTTAAGCTGAGAGGATGGCAGTAGAAGAAAAATGAGTTATCA 424  
Db 8906 AAGTAACATTTAAATAATAACAATAAAGCATATGAGTTAATTAATAAATGATGTTA 8847  
Qy 425 ATGCTGTTCTTTCAATCTATTAATTCACAAAAGAGATATTTAAGAAAGCTTTGAGAG 484  
Db 8846 TTTTGAATATAAATGTTTCAACAGCATATGAATTCACGACACAAAATTTGAACATA 8787  
Qy 485 AAGCAAAACATGTAAGTTTAAATGAAATCATTCAGTAGATGAGAAAGTAAATTC 544  
Db 8786 AAGTTATGATGATGATGTCACAGCATCAACAAATTAACACACAAAGAAATTAAGTTA 8727

Qy 545 CATTGAATCCGAATGAGACATTAACGTTAGAGAGAAAAATCAAT---GCTGTTGAGGCA 601  
Db 8726 AAATGATGATATTCAGAGAGCTTGAATATAAATAAGCTGAATGTTATGTAATCAAA 8667  
Qy 602 TCGGTTTATATGCGGCGGATTTAGATTTGAAGATCTGCATATCTAAAGCAGAAATA 661  
Db 8666 ACATTTTAAAGGATTCGAAATTAATAACACAGTGAATTAATACAAATATGTTGATC 8607  
Qy 662 CAGATTTTAAAAATTTAGTCAATATTAAGTATCAATTAATTTGCTGTGACCGGAAT 721  
Db 8606 TAAATTTGATTAATAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8547  
Qy 722 TAAAGCTACCAAGACAAATCTGAGATATTAATTTCTTACGTACATAGATTTCTCTC 781  
Db 8546 ATAGTTTAAAGAAATTAATTAATTAATTAAGAGTTGACAAATGCGCAAGATCAGAAATTTGTA 8487  
Qy 782 AAAAAGCTATGAGAAAAATTTCACTGTTGAAAGAGAAATAGAAATATGTAAAGGA 841  
Db 8486 AAGAAATCAATGTAAATTAATTCATTTGATTTAATTTAACAAGGAAATTAAGCATCAA 8427  
Qy 842 ATACCAAGCAAAATATGAATCTGATGCTGATTTGGAAGCAGATGAAATTAATAATTA 901  
Db 8426 GCGTAGAAGAAATTAATGATGCTGCTCTGATTAATTAATTAATTAATTAATTAATTA 8367  
Qy 902 GTGCAAGCTCAAAATGAGAGATTTAATAAGAAAGAGGAAAAAGAACTTTAACA 961  
Db 8366 ATTTAAAGATGACAAATGATTTTAAACAAATGATTTGCAACAACTTACTATGATA 8307  
Qy 962 CTCCTTAAGTTATCAATGTGAGAGCTTCGTAAGAGTAAATTAAGAAAAATCATAG 1021  
Db 8306 ACGAACAAAAGGTGATGCTATTTGTAATAACTTAATGACAAAATCAAAAATTTAAGAG 8247  
Qy 1022 GAAAGATGTTGACATTAACAGCTGAGCAAAAAGATTTCTATGATGCAACTTATGACTA 1081  
Db 8246 CAATATTAACAAATTTGATTTAATAAATGTAATTAATTAATTAATTAATTAATTA 8187  
Qy 1082 AGCTTGAAAGCACTCTTTTAAAGCTTTGTAACAGTTCTAATTTCTCATCAATTAATG 1141  
Db 8186 ATTTATCTCAAACTTTATTAATTAAGTGGCAAAAATTAATTAACAAATGTAATTAAGCA 8127  
Qy 1142 GATTTTAAAGTTTATTTGACAAAGTAA 1166  
Db 8126 GTAATTTAAATTAATTAATTAATTA 8102

## RESULT 32

US-08-728-323A-1

Sequence 1, Application US/08728323A

Patent No. 5948676

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James J.

APPLICANT: Edelman, Isidore S.

APPLICANT: Moore, Patrick S.

TITLE OF INVENTION: Immediate Early Protein From Kaposi's

TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

TITLE OF INVENTION: Encoding Same And Uses Thereof

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/728,323A

```

?      FILING DATE:
?      CLASSIFICATION: 435
?      ATTORNEY/AGENT INFORMATION:
?      NAME: White, John P.
?      REGISTRATION NUMBER: 28,678
?      REFERENCE/DOCKET NUMBER: 05
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: 212-278-0400
?      TELEFAX: 212-991-0525
?      INFORMATION FOR SEQ ID NO: 1:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 3469 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: DNA (genomic)
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 1..3469
?      OS-08-728-323a-1

```

QY 7989 ATCTCAGAGAGCTTCTGTAGAGCAGTACAGGCGCAGGAATTTATTTCCGAAGCAAAAGA 8048  
DB 2387 AGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGAGC 2446  
QY 8049 TGC CGAAGCTCTTATTGAAAGTTAGTACAAATCCCGAAGAACTATTTTTCATGCGA 8108  
DB 2447 AGGAGTTAGAGAGCAGAGCAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGG 2506  
QY 8109 TAAATGTAATGAGCAACACATATAATGAAATGAAATGAAATGAAATGAAATGAAATG 8168  
DB 2507 AGCAGAGTTAGAGAGCAGAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGC 2566  
QY 8169 AGCTTCTGTATTGAGAGAGTGTAGTCAACAGGCGAAGCTACTGCTGAGTAAAC 8228  
DB 2567 AGGAGCAGAGAGCAGAGAGTTAGAGAGAGTTAGAGAGCAGAGCAGAGCAGAGAGC 2626  
QY 8229 TATGTTAGAGTTGAGAGAGAAATTTGTTCAAGAACAAATGCAATTCATTTCTTA 8288  
DB 2627 AGGAGAGCAGAGAGTTAGAGAGAGTTAGAGAGCAGAGAGAGAGTTAGAGAGAGTGG 2686  
QY 8289 AGTAAAGAGTTGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 8348  
DB 2687 AAGAGCAGAGAGAGCAGAGAGTTAGAGAGAGTTAGAGAGCAGAGAGCAGAGAGTTAGAG 2746  
QY 8349 AGGAGGAAATGCGGAGCAGAGAGTGAATCTTCTACAGCAGCAAAAGTAACTGAA 8403  
DB 2747 AGGTGAAGAGAGCAGAGCAGAGGAGGTGAACAGAGAGCAGAGAGCAGAGAGCAGTGA 2801

## RESULT 35

US-09-894-273-1  
Sequence 1, Application US/09894273  
Patent No. 6756203  
GENERAL INFORMATION:  
APPLICANT: Kieff, Elliott D.  
APPLICANT: Ballesteras, Mary E.  
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
TITLE OF INVENTION: VIRUS DNA TO MEDIANTE EFFICIENT EPISOME PERSISTENCE  
FILE REFERENCE: 16412-10001R  
CURRENT APPLICATION NUMBER: US/09/894,273  
CURRENT FILING DATE: 2001-06-28  
PRIOR APPLICATION NUMBER: US 60/109,422  
PRIOR FILING DATE: 1998-11-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 3489  
TYPE: DNA  
ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-894-273-1

Query Match 0.6%; Score 59; DB 3; Length 3489;  
Best Local Similarity 45.3%; Pred. No. 0.034;

Matches 215; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 7929 TAAAGGAATTGGAAGAAATCAATGTCAGAGTGAAGAAAAGAGAGTACTGCTGA 7988  
DB 2327 TAGAGAGATCAGAGCAGAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGCAGG 2386  
QY 7989 ATCTCAGAGAGCTTCTGTAGAGCAGTACAGAGGCGCAGGAATTTATTTCCGAAGCAAAAGA 8048  
DB 2387 AGTTAGAGAGCAGAGCAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGAGC 2446  
QY 8049 TGC CGAAGCTCTTATTGAAAGTTAGTACAAATCCCGAAGAACTATTTTTCATGCGA 8108  
DB 2447 AGGAGTTAGAGAGCAGAGCAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGG 2506  
QY 8109 TAAATGTAATGAGCAACACATATAATGAAATGAAATGAAATGAAATGAAATGAAATG 8168  
DB 2507 AGCAGAGTTAGAGAGCAGAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGC 2566  
QY 8169 AGCTTCTGTATTGAGAGAGTGTAGTCAACAGGCGAAGCTACTGCTGAGTAAAC 8228

DB 2567 AGGAGCAGAGAGCAGAGAAATTTAGAGAGGTTGAGAGAGCAAGCAGAGCAGAGAGCAGAGC 2626  
QY 8229 TATGTTAGAGTTGAGAGAGAAATTTGTTCAAGAACAAATGATGAAATGCAATTTCTAA 8288  
DB 2627 AGGAGAGCAGAGAGTTAGAGAGGTTGAGAGAGCAGAGAGCAGAGAGTTAGAGAGGTTG 2686  
QY 8289 AGTAAAGAGTTGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 8348  
DB 2687 AAGAGCAGAGAGCAGAGAGTTAGAGAGAGTTAGAGAGCAGAGCAGAGAGCAGAGAGTGAAG 2746  
QY 8349 AGGAGGAAATGCGGAGCAGAGAGTGAATCTTCTACAGCAGCAAAAGTAACTGAA 8403  
DB 2747 AGGTGAAGAGAGCAGAGCAGAGGAGGTGAACAGAGCAGAGCAGAGAGCAGAGAGCAGTGA 2801

## RESULT 36

US-08-770-379-20/c  
Sequence 20, Application US/08770379  
Patent No. 5849564  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/770,379  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 52342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-770-379-20

Query Match 0.6%; Score 59; DB 2; Length 3207;  
Best Local Similarity 45.3%; Pred. No. 0.07;

Matches 215; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

QY 7929 TAAAGGAATTGGAAGAAATCAATGTCAGAGTGAAGAAAAGAGAGTACTGCTGA 7988  
DB 19670 TAGAGAGATCAGAGCAGAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGCAGG 19611  
QY 7989 ATCTCAGAGAGCTTCTGTAGAGCAGTACAGAGGCGCAGGAATTTATTTCCGAAGCAAAAGA 8048  
DB 19610 AGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGCAGAGAGTTAGAGAGCAGAGAGC 19551

Qy	8049	TCGCCGAAGCTCTTATTTGAAAGTAGTACAAAAATCCGAGAAAGTATTTTTCATGCA	8108
Db	19550	AGGAGTTAGGAGCAGCAGACAGAGATTAAAGAGCAGGAGCGAGTTAGAGAGCAAG	19491
Qy	8109	TAATGTAATATGGAACCAACACTAAATGAAAGTAAACGACGTTCTTAAAGCACTAAC	8168
Db	19490	AGCAGAGTTTAGAGAGCAGAGAGTGGAAAGCAGAGCGAGAGGTGGAGAACCAAGAC	19433
Qy	8169	AGGTTCTGTATTTGGAGAGGTTGGAGTCACCAAGCGCAGAAAGCTACTGCTGCAGGTAAAC	8228
Db	19430	AGGAGCAGGAAGACAGCAAGATTAGAGAGGTGGAGAGCAGCAAGCAGAGCAGAGAGAC	19371
Qy	8229	TATGTTAGAAATTGGAGAGCAAAATTTGTCAGAACAAATCGATTGAATGCAATTTCTAA	8288
Db	19370	AGGAGGAGCAGGAGTTTAGAGAGGTGAAGAGCAGGAAGCAGGAGTTAGAGAGGTGG	19311
Qy	8289	AGTAGAAGTTTGGATGAGTAAAGTAACTGCTAAATCTTCTGTAGTATCAGAAATGG	8348
Db	19310	AAGAGCAGGAAGACAGACGAGTTTAGAGAGGTGGAGAGCAGGAGCAGCAGAGTTTAGAG	19251
Qy	8349	AGGAGCAATTTCCCGAGCAGAGAGTAATCTTCTACAGCAGAAAGTAAATCTGAA	8403
Db	19250	AGGTGGAAGACAGACAGCAGCAGGAGGTGGAAACAGCAGGAGCAGAGACGCTGGA	19196

```

RESULT 37
US-08-757-669A-20/C
/ Sequence 20, Application US/08757669A
/ Patent No. 6183751
/ GENERAL INFORMATION:
/ APPLICANT: Chang, Yuan
/ APPLICANT: Bohenzky, Roy A.
/ APPLICANT: Russo, James J.
/ APPLICANT: Edelman, Isidore S.
/ APPLICANT: Moore, Patrick S.
/ TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
/ TITLE OF INVENTION: SEQUENCES AND USES THEREOF
/ NUMBER OF SEQUENCES: 20
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Cooper & Dunham LLP
/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/757,669A
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28, 678
/ REFERENCE/DOCKET NUMBER: 45185-F
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 391-0525
/ TELEFAX: (212) 278-0400
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32207 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-08-757-669A-20

Query Match 0.6%; Score 59; DB 3; Length 32207;
Best Local Similarity 45.3%; Pred. No. 0.07;
Matches 215; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

```

QY	7929	TAAAGGAATTTGAAAGGAAATCAATCTCAAACTCGAATAAAGAAAACAAGATGACTCGCA	7988
Db	19670	TAAAGGATCAGGACGAGGATTAGAGAGACAGAGACAGAGAGTTAAGAGAGCAGAGCAGG	19611
QY	7989	ATCTCAAGGAGCTTCTGTAGAGACAGTAGCAGGCGCAGAAATTATTTCCGAGCAAAAGA	8048
Db	19610	AGTTAAGAGAGCAGAGAGCAGGAGTTTAGAGAGCAGAGAGCAGAGAGTTAGAGGACGAGAGC	19551
QY	8049	TGCGCGAGCTCTTATTTGAAAAGTTAGTACAAATTCGGAAGAGATTTTTTCATGCAGA	8108
Db	19550	AGGAGTTAGGAGACAGAGCAGAGCTTAGAGAGCAGAGAGCAGAGATTAGAGAGCAGG	19491
QY	8109	TAATGTGAATATGGAAGCAACACATPAAATGAAAGTAACAGCAGATTTCTTAAAGCAGTAAC	8168
Db	19490	AGCAGAGCTTAGAGAGACGAGAGCTGGAAGAGACAGACAGAGAGGTGGAAGAGCAGAGC	19431
QY	8169	AGGTTCTGTATTGAGAGAGGTTGGAGTCCAAAGGCAGAGCTACTGCTGCAGTTAAAC	8228
Db	19430	AGGAGCAGAAAGCAGGAATTTAGAGAGGTGAGAGCAGACAGAGCAGAGCAGGAGCAGC	19371
QY	8229	TATGTGTGAATGTGAGAAAGAAATTTGTCAGAACCAATGCATGGAATGCAATTTCTAA	8288
Db	19370	AGGAGGACGAGGATTAGAGAGCTGGAAGACAGAGAGCAGAGGATTAGAGAGGTGG	19311
QY	8289	AGTAGAAGTTGTGATGATGAATPAAATCTGCTAAATCTTCTGTAGTATCAGAAATGCG	8348
Db	19310	AAGAGCAGAAAGACAGAGGATTAGAGAGGTGGAAGAGCAGAGCAGCAGGAGTTAAGAGG	19251
QY	8349	AGGAGGAATTTCCGAGACGAGAGTGAATACTTCTACAGCAAAAGTAATCTGGA	8403
Db	19250	AGGTGGAAAGCAGAGACAGCAGAGGGGTGGAACAGCAGGAGCAGAGACGGTGGGA	19196

```

RESULT 38
US-09-230-371A-20/c
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; TITLE OF INVENTION: USES THEROPF
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230, 371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20

Query Match          0.6%; Score 59; DB 3; Length 32207;
Best Local Similarity 45.3%; Pred. No. 0.07;
Matches 215; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

Qy      7929  TAAAGGAATTGGAAGAAATCAATGCAAGTGGAAAAAGAAAACAGAGTGACTGCTGA 7988
Db      19670  TAGAGGATTCAGGACAGAGAGTTAGAGGACGAGCAGAGCTTTAGAGGAGCGAGACAGG 19611|||||
Qy      7989  ATCTCAAGSAGCTTCTGTAGAGCAGTACGAGGCGCAGGAGATTTATTTCCGAAGCAAAAAGA 8048
Db      19610  AGTTAGAGGAGCAGACGACAGAGAGTTTAGAGGACGAGAGCGAGAGTTAGAGGAGCAGAGAGC 19511|||||
Qy      8049  TGCCGGAAGCTCTTATTATTGAAAGTTAGTACAAAATCCGAGAGAGAGTATTTTTCATGCGACA 8108
Db      19611  TGGGAGGAGCTCTTATTATTGAAAGTTAGTACAAAATCCGAGAGAGAGTATTTTTCATGCGACA 19511|||||

```

Db 19550 AGGAGTTAGAGGAGCAGAGCAGAGTTAGAGAGCAGAGGAGTTAGAGAGCAGG 19491  
Qy 8109 TAAATGTAATATGAGAGCAACATATAATGAAATACAGAGTTCTTAAAGCAGTAC 8168  
Db 19490 AGCAGAGTTAGAGAGCAGAGAGTGTGAAGAGAGCAGAGAGGTTGAGAGCAGAGC 19431  
Qy 8169 AGGTTCTGATTTGAGAGAGTTGAGATCCAAAGCAGAGAGTCTGTCAGGTAAAC 8228  
Db 19430 AGGACAGAGAGAGCAGAGATTAGAGAGGTGAGAGAGCAGAGAGCAGAGAGC 19371  
Qy 8229 TATGTAGAGAGTTGAGAGAGAAATTTGTCAGAAACAATCGATTGAATGCAATTTCA 8288  
Db 19370 AGGAGAGAGAGAGTTAGAGAGGTGAGAGAGCAGAGAGAGAGAGAGAGTTAGAGAGGTGG 19311  
Qy 8289 AGTAAAGAGTTGATGAGAGATTAAGTAACTGCTAAATCTTGTATGATCAGAAATGG 8348  
Db 19310 AAGAGCAGAGAGAGCAGAGAGTTAGAGAGGTGAGAGAGCAGAGAGCAGAGTTAGAGG 19251  
Qy 8349 AGGAGGAATTCGCGAGCAGAGAGTAACTTCTACAGACCAAGTAACTAGAA 8403  
Db 19250 AGGTGAGAGAGCAGAGCAGAGGAGGTGAGAGCAGAGAGAGAGAGAGAGCAGTGA 19196

RESULT 39  
US-08-682-517-13  
/ Sequence 13, Application US/08682517  
/ Patent No. 5874267  
/ GENERAL INFORMATION:  
/ APPLICANT:  
/ TITLE OF INVENTION: Expression of surface layer proteins  
/ NUMBER OF SEQUENCES: 25  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ OPERATING SYSTEM: IBM PC compatible  
/ SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
/ CURRENT APPLICATION DATA:  
/ APPLICATION NUMBER: US/08/682,517  
/ FILING DATE:  
/ CLASSIFICATION:  
/ INFORMATION FOR SEQ ID NO: 13:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 3666 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: DNA (genomic)  
/ US-08-682-517-13

Query Match 0.6%; Score 58.2; DB 2; Length 3666;  
Best Local Similarity 41.7%; Pred. No. 0.05;  
Matches 801; Conservative 0; Mismatches 1093; Indels 27; Gaps 6;

Qy 5830 ATAGCAGAAATTAATCATGCAAAATCACTCGGAAGGAAATGTCGAGATTATTAAGAG 5889  
Db 1732 AAAAGAAAACAATCTCTGTAATAAACAATGATTAAGGTAAGTCAATTTGTATCGGT 1791  
Qy 5890 TCTGATGCGTAAATGCTTAATATGAGAAACAGTGTCTGAGTGGCCCGTGCAGCAATA 5949  
Db 1792 TCTGATACGTAAGATTAATGCAACCAATTTGCTGATGATTAATTAATCTTCTGAT 1851  
Qy 5950 GGAGGCTCAACAGTGTGAATGAATTAAGAGATCTCAAAAGCAATATGTAAGAGTTCT 6009  
Db 1852 GCAAAACAAGGAGCACTTATGATGAGTGAACCAAAAGCAGTTGCACCAATCTCTTACTTC 1911  
Qy 6010 ACAGTATTTGCTAAAGAGAAAACAGATGATTAATTAATTAATCAAGGCAAGTATGATAA 6069  
Db 1912 CAAGCAGCATATCTGATGCTCAGCTATCAAGCATACAAAAATCAGATCTTATATAA 1971  
Qy 6070 GTGTGATGATTAAGTATTCAAAAATCTTAATATTAAGAGAACTTATCAGAAAAAGAAA 6129  
Db 1972 GCTGTAACTAAAGTTGATGATTTCTGAAACGTCAGATATTTTGCAGCAGAAATTAAGTAAACCA 2031

Qy 6130 ATAAATATATAAAAAAGATTTGTTACCAATAGTTCAGTACTCATCTTAAAACTTTA 6189  
Db 2032 AGCGGCAAAAAAGTAACTGATCTTATTAAGAAAGCACTTATTAACAATCTCAACTACT 2091  
Qy 6190 TTGGCAATGCGCTGCTGTCAGACAGCCGAGTGGCAGAACTGTTAATATCAACAG 6249  
Db 2092 GGTCTATATGATTAATTAAGTATGATTAACCAATTTATCTCAACAAATCGTAGCTACACAGTA 2151  
Qy 6250 GTTTATGAGAAACAGAAAGCTCTGTGAAAAATCTATTAATTAATGCAAAACATTATCT 6309  
Db 2152 ACTTATGAGACTCTTATTTCTTACAGAACTGTTATTAACCTGCTGAAGAAATTTAGAA 2211  
Qy 6310 GTAAATCAGAGATTAACCAATTCATCCGAGTATGAGTCTGTGTTGTGTTGGA 6369  
Db 2212 GTTACTTACAGTATGATTAAGAAACAACTGCTTAAAGTAATTTGCTACAGATATGCTGT 2271  
Qy 6370 AATGAGAGTATGAGAGCTTCTTGTATTCATATTTAATAAGAAATCCAGACAGAA 6429  
Db 2272 AATACAGACGTTAAAGACTATGCAATTTACTGCTAAAGAGCTACAGCTACATTCACAGCT 2331  
Qy 6430 GTTGAAAAAACTCAATGCTGATGAAAGTTTCGAGAGAAAGCTGAAATTTACAGCAGAT 6489  
Db 2332 ACAAATGAAAGTTCCAACTCTTACACTGCTGATGACTCAAA--TTCAATACAGCTGAT 2388  
Qy 6490 TCTAAGCAGAGAAATTTCTCTTTTGAAGTCCGAGTCCGAGAGCCGAGTCCGAGAGC 6549  
Db 2389 TCTGTTCAACACCACTATTTGTTGTTGCTGTGTAAGAAACCAAGTAAATATGCTGT 2448  
Qy 6550 GTGCGAGAACCGTTCCGTAATCAATTTGACAGAAAGCAGAGAGTATGATGAGAGAA 6609  
Db 2449 GTATCAGCAGCAAAATATAATTAATCTTCGAGCTAATGTAATGAAATTTTGAGAAACG 2508  
Qy 6610 GCA----AAGATTTTGTGTAATAAAGCTGATTAACAGCAAAACGTTAATTCCTGTTGC 6665  
Db 2509 GCATGAGGAGATTAATTAATCAATATGCAATGCACTGAGGCAAAAGTAAATTCATAT 2568  
Qy 6666 AA--TTGGAATGCGCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCC 6723  
Db 2569 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2628  
Qy 6724 GCAGTTACCAAGATGATTAATCAACACAGAGCAGAGAGTGAATAAT--TCTAAATTA 6778  
Db 2629 GCTATCAACAGAGTGTGCTCAACCAACACAGAGTCCACATGAGGCGCATTAACATTA 2688  
Qy 6779 TGACTCGAAGAGT-----TAGATGATTAAGAGAAATGATTAATTCAGGTACT 6831  
Db 2689 ACACGAGAGCTGTGTTGATGATTAATCACTGCACTTAACATTTGAATTTCA 2748  
Qy 6832 GAAATCGGTTGAGCGAGAGTGAATTTCTGAGCGGAGTATCTGAGTGTGCTGTC 6891  
Db 2749 TTAGCTGATGAGATCTTAATGTAAGTCAACCACTGTTGATGCAACTGTTTCAATTA 2808  
Qy 6892 AATAATATTGCAAAATAGATGAAACAGATATCGATATGATCTTACCTTTCTACT 6951  
Db 2809 AAAAGATGCAAAATTAATCTTATCTCTTACATTAAGTGAACGCTGCTAATACAGGT 2868  
Qy 6952 GATGTAATGTAAGAGCTCTTAATTAATTTGCAATTCCTGACAGCCGCTGAGAGAGCC 7011  
Db 2869 GTATTTGCTACAACTGTTCAAGCTGTGATCATTAATCTTTTAACTGCTGTGATCAATTA 2928  
Qy 7012 GCAGGCTTTCGAGAGATTAACGAGAGTGTCTTCTGTTACATTAATTAATGTTCTGTGATA 7071  
Db 2929 GTTACTTATGAGATGCTTAATAATGCTGAGAGTGTGCTGAAAAATTAATGCTGAGCGTA 2988  
Qy 7072 GCTGAGTTCACAATTAATCTGATTTGATTTCCGTACGAGAAAAAGTAATGTAAGGCA 7131  
Db 2989 ACATTAAGAAAACTACTGAGAGCAATTAATCTGATACATTAACAGAGTATTAACCA 3048  
Qy 7132 AAAAGGAAAAAATATTAAGCAACACAGCAAAATGAGAGATTCGAGAGCAGCAATC 7191  
Db 3049 TCAGAGCTTACAGAGCTGAAATATATCTTC-----TAATCAATTTGCTGAGATTAATCA 3102  
Qy 7192 GAGGCAATGTCTTGTAATTAATTTTGAACAGCTGTAGAGATGAGAAAAATTTCTGAA 7251

Db 3103 TTTCACACAGGTGAGGATTCACCTTTAAATATGATTAATGCTGCTCAAGTAATTAC 3162  
Qy 7252 GGAAGAAGACAGAGTTTAAAACTTTAGACGAGTAACAAAGAACAGATTAATAA 7311  
Db 3163 TTACGAGGTAAAAAGGTGACAGAGGTAGCTAGCTATCAATGCTACATTTGCAAGT 3222  
Qy 7312 GTAATGATGCTACGAAAAAATCTTACAAATCAGCAGGTATTTTACAGAAATCTTCT 7371  
Db 3223 ACTGCACTGTTTCTGAGACAAAGTAGTTAATTAATCAGCTACAAACAGGTGGTTCT 3282  
Qy 7372 GTAAAAAGCGATAGAGAGATCTCAGGAGAGAAATTAAGCATTTGTGAAGCTTCT 7431  
Db 3283 GAAGTTGAAGTTACATCTCTCTGTTAATCAAGATTAATAAGCAGTAAACGTTAA 3342  
Qy 7432 GATATTTATGAAAAATGATGATTTAACAAGAGACAAAGATTAATCAGCTTCTACT 7491  
Db 3343 GATCAAGTCGTTGACAGAACAGCTCTCAAAAAGCATTCAGATTAATCTACAGCCCTTCT 3402  
Qy 7492 GGTGTTTGGGAATGACAGGTCTTTCCTCCGATCAGAAACAGTGCAGTTACAAATAT 7551  
Db 3403 GTGGGTGAAAAAGTAGTTATGATGATGTAATATGCTGTAGCATTTGGAACTGCT 3462  
Qy 7552 AAAAGAAATTCGGAGTTACTCTTTGAAATTTCTTTTGGAAAGCAGCTGAAAAATTAAT 7611  
Db 3463 CCAACAGCAAAATACATTCGTAGTGAATCTGCTGCTAAATACATTAAGCTTACAGTAC 3522  
Qy 7612 GTTAGATGGAATATTAACAGAAATGTGCTTAAACAGCAATCAAGGCTCTGTAGAGAGA 7671  
Db 3523 CAAGTGCAAATCTTGCTGCTACAAATGATCTTTAAACATGCGATTAAGTTACAGCT 3582  
Qy 7672 TTGGGAATAGAGCTGCTATGACAGAAATTAATCTAAATGAGATCAAAATTCAGTAT 7731  
Db 3583 TCTGCAACAGGTCTACTATTATTAATTAATCTTACTGTAACACCAAGTGTGATCAAT 3642  
Qy 7732 A 7732  
Db 3643 A 3643

RESULT 40  
US-08-682-517-14  
; Sequence 14, Application US/08682517  
; Patent No. 5874267  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Expression of surface layer proteins  
; NUMBER OF SEQUENCES: 25  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/682,517  
; FILING DATE:  
; CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3666 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULAR TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..3666  
US-08-682-517-14

Query Match 0.5%; Score 58.2; DB 2; Length 3666;  
Best Local Similarity 41.7%; Pred. No. 0.05;  
Matches 801; Conservative 0; Mismatches 1093; Indels 27; Gaps 6;

Qy 5830 ATAGCAAAATTAATCATGCAAAAAATCAGTCGGAAGGAATGTCGAGTATTACAGAG 5889  
Db 1732 AAAGCAAAACAAATCTCTGTAATAACAAATGATTAAGGTAGGTACTATTGTTGCGT 1791  
Qy 5890 TCTGATCCGTAATTTGCTAATTTATGACAGAACAGTGTCTGAGTGGCCGTCAGCAATA 5949  
Db 1792 TCTGATACAGTAACGATTAATGCAACCAATTCCTTGATGATTAATTAATCTTCTGAT 1851  
Qy 5950 GGAAGCTCAACCAAGTGAATGAATAATACAGATCTACAAAGCATATGTAAGAAATCT 6009  
Db 1852 GCAAAACAAAGGCGACTTGAATGAAGGTGAACCAAAAGCAGTGAACCAATCTTACTTC 1911  
Qy 6010 ACAGTATGCTAAGAAAGAAACAGATGATTAATTACTACTAAGGCGAATGATTA 6069  
Db 1912 CAAGCACCATTCTTGATGCTCAGCTATCAAGCATACAAAAATCAGATCTTAATTA 1971  
Qy 6070 GTGATGATTAAGTATCAAAAATCTTAATTTAAGCAAGACTTATCAAAAAAGAAA 6129  
Db 1972 GCTGTAATTAAGTTGATGCTTGAAGCTGCAAGATTTGACAGCAATTTAGTAACCA 2031  
Qy 6130 ATAAAGTAATAAAAAAGATTTGTTACCAATAGTTCAGTACTCATCTTTAAATCTTTA 6189  
Db 2032 AGCGCAAAAAAGTAAGTACTGTAATTAAGAAAGCACTTATTAATCAATCTACAAATCT 2091  
Qy 6190 TTGCAAAATGCCGCTGCTTCAAGCAAGCCGAGTGCAGGAACGTTAATATCAACAAG 6249  
Db 2092 GGTGTAATGATTAATTAAGTATTAACCAATTAATCTCAACCAATGTAAGTCAACAGTA 2151  
Qy 6250 GTTATGAGAAACAGAGCTCTGTGAGAAAATCTATTAATTAATGCAAAAATTAATCT 6309  
Db 2152 ACTTATGAAGCTACTTTATCTTACAGGAACGTGTTATTAACCTGCTAAGATTTAGA 2211  
Qy 6310 GTAATAATCAGAGATTAACAGAAATTCAGAGTAGAGTTCGTTGTTGTTGTTGTTG 6369  
Db 2212 GTTACTTCAAGGATGTAATAACCACTGCTGTAAGTAATGCTACAGTATTCCTGTT 2271  
Qy 6370 AATGAGAGATGAGAGCTTCTTGATACCAATATTAATTAATAAGAAATCAAGCAAGA 6429  
Db 2272 AATACAGACGTAAGAAAGATATGATTAATCTGCTAAGAAAGCTACATTCACAGCT 2331  
Qy 6430 GTTGAATAATCTACATGCTGATGAGATTTGAGAAAGAGTGAATTCACAGCAT 6489  
Db 2332 ACAATGAAGTTCAAATCTTTACACTGTGTGATCTACAA---TTCAATACAGTGT 2388  
Qy 6490 TCTAAGCAAGAAATTCCTCTTTTGGAGTCGAGTCGACAGCCGGGTAGAGCCGGA 6549  
Db 2389 TCTGTTCAACACAGCACTATTTGTTGCTGTAATAAACCCAGTGAATATCTGCT 2448  
Qy 6550 GTGCAAGAACCGTTTCCGTAAATCAATTTGACAGAAAGACGAGTGAATGTGAAGA 6609  
Db 2449 GTATCAGGCAAAACATTAATAATCTCGAGCTAAATGTAAGAAATTTGTAAGCG 2508  
Qy 6610 GCA---AAGATTTTGTGTAATAAAAGCTGAGATTCAGCAAAACGTTATAGTTCTGTTGC 6665  
Db 2509 GATGGAAGGATTTAATTAATCTAATATGCACTGAAGGCCAAAGATTAATCATCTCAT 2568  
Qy 6666 AA--TTGGAATGCCGAGTCGAGTCGAGTCGAAAGAGCTGAATGAGCAGCAGTG 6723  
Db 2569 AATGTAGATGATGATACAGTTACATTTAAGTAATTAAGTGTGTTAATCTTCAACTGA 2628  
Qy 6724 GCAGTTACCAAGATGAATCAAAACGAGAGCAAGATCAAAAT---TCTAAATTA 6778  
Db 2629 GCTATCAACACAGTTGCTCCAAACACACAGAGCTCCAACTACGCGCATTAACATTA 2688  
Qy 6779 TGACTGAAACAGT-----TAGATGTAATGAGAAATGAGTAATAATCAGGTACT 6831  
Db 2689 ACACAGACAGCTGTGTGTTTATGTTTAACTGCACTACATCACTTAGAATTTGA 2748  
Qy 6832 GGAATCGTTACCGGAGCTGAATTTCTTGACGCGAGATCTGAGATGTTCTGTC 6891  
Db 2749 TTAGCTGATGACATCTTAATGTAAGTGCAACAGTGTGATGATCACTGTTCAATTA 2808  
Qy 6892 AATTAATTTGCAAAATGAGTGAAGAAACAGATATGATCATTAATCTTCTACT 6951

Db 2809 AAAGATAGTCAAAATTAATTCATTATCTTACATTAGTGAACCTGGTCTAATACAGGT 2868  
QY 6952 GATGTAATGTAAAGCTCTTAATAAAATTCGAATTCCTTGACAGCCGGTGAAGAGCC 7011  
Db 2869 GTATTGTCTACACACTGTTCAAGCTGGTACATTAATCTCTTAACCTGCTGACATTAAACA 2928  
QY 7012 GCAGGCTCTTGACAGCTTACCCGAGTGGTTCTGTTAACACTATTAATAGTTCTGTGATA 7071  
Db 2929 GTTACTTATGCAAGATGCTAAAAATGCTGAGGTTGCTGAAAAATATTACTGCTAGCGTA 2988  
QY 7072 GCTCGAGTTACAAATTACTCTATTTGACTTCCGTACGAGAAAAAGTAAATGTAACGGCA 7131  
Db 2989 ACATTAAAGAAAACCTACGAGCAATTACTTCTGATACATTTACACAAGGTGATTACCA 3048  
QY 7132 AAAAGGAAAAAAATATTAAAGCAACAGCACAATGCAAGAAATCGAAGAGCAGACAATC 7191  
Db 3049 TCAGCAGCTACACGACGCTGAATATACCTC-----TAAATCAATGCTGCAGATTATACA 3102  
QY 7192 GGAGCCAAATGCTTGGTAATTAATTTTGACAGCTGTAGAAAGATGAAGAAAAATTCGAA 7251  
Db 3103 TTTGCAACAGGTGAAGGATTCATTTAAATATTGATTAATGCTGCTGCAAGTAAATTAAC 3162  
QY 7252 GAAAAAGGAAAGGATTTTAAAACTTTAGACGAAGTTAACAAAGAACAGATAAAAA 7311  
Db 3163 TTAGCAGGTAAAAAAGGTGCACAAGGTGACGATGCTATCAATGCTACATTTGCAGGT 3222  
QY 7312 GTAATGATGCTACGAAAAAAATCTTACAAATCAGCAGGATTTCTACAGAAAGTACTTCT 7371  
Db 3223 ACTGCACCTGTTCTGGAACAAAGTAGTTAATAATCAGCTACAAACAGGTGTTGTTCT 3282  
QY 7372 GTAAAAAGCGATAGAGAGATCTCAGGAGAGAGAAATTAAGCCATTGTGAAGACTTCT 7431  
Db 3283 GAAAGTTGAAGTACATTCCTCTGTTAATCAAGTATTAAATGCAAGTATTAAACGTA 3342  
QY 7432 GATATTATTGAAAAAATGTAGATTTTACACAGAGACAGAGAAATTAATCACTTCTACT 7491  
Db 3343 GATCAAGTCTTGACAGAAACAGCTCTACAAAAGCATTCACGATTACTACAGCCCTTCT 3402  
QY 7492 GGTGTTTGGGAACTGCAGGCTTCTCCGATCAGAAACAGTGCAGTTACAAATATT 7551  
Db 3403 GTGGGTGAAAAAGTATTGTTATGATGTTGTAATATACTGCTGTAGCAATTTGAACTGCT 3462  
QY 7552 AAAAGAAATTCGAGATTACTGTTGAAAAATCTTTTGGAAGACAGCTGAAGAAATTAAT 7611  
Db 3463 CCAACAGCAAAATACATTCGTAGTTGATCTGCTGAATACATTAAGCTTCAGTAGCTGAC 3522  
QY 7612 GTTAATCGAATATTACAGGAAATGTGCTTTAACAGCATATCAAGGTCGTAGAGCA 7671  
Db 3523 CAAGCTGCAAATCTTGCTGCTCAATGTATCTTTAAACACTGCAGATAAGTTTACAGCT 3582  
QY 7672 TTGGGAATAGAGCTGCTATGAGAAATTAATTTCTAATGGAAGATCAAAATATCAGTATT 7731  
Db 3583 TCTGCAACAGGTGCTACTATTACATTAACTTACTGTACACACAGTAGTACTAACAAAT 3642  
QY 7732 A 7732  
Db 3643 A 3643

Search completed: November 25, 2005, 20:49:28  
Job time : 1096 secs

20-2

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: November 25, 2005, 03:41:50 ; Search time 30972 Seconds  
(without alignments)

17850.297 Million cell updates/sec

Title: US-10-647-057-8

Perfect score: 9726

Sequence: 1 atgagcggcgcacacaaataa.....gagaaaaagggaataatga 9726

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_rm:\*  
5: gb\_ov:\*  
6: gb\_par:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_hcg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9726	100.0	9726	6	AR442020 Sequence
2	9726	100.0	11130	6	AR442027 Sequence
3	9726	100.0	12700	1	AF312861 Fusobacte
4	2780	28.6	2780	6	AR442022 Sequence
5	2322	23.9	2322	6	AR442025 Sequence
6	2141	22.0	2141	6	AR442023 Sequence
7	1887	19.4	1887	6	AR442024 Sequence
8	1130	11.6	1130	6	AR442021 Sequence
9	131.6	1.4	170221	9	AC127421 Mus muscu
10	130.2	1.3	179553	14	AC024253
11	123	1.3	4995	2	PFAMESA
12	121	1.2	271546	2	AE014843 Plasmodiu
13	120.4	1.2	144745	14	AC136928 Homo sapi
14	120.2	1.2	38662	14	AC138477 Homo sapi
15	119.8	1.2	36348	14	AC138510 Homo sapi
16	119.4	1.2	347050	2	PFAS29351 Plasmodiu
17	118.8	1.2	5420	2	AF056936 Plasmodiu
18	116.2	1.2	5361	6	AS8933 Sequence 2

19	116.2	1.2	5361	6	AR131272	AR131272 Sequence
20	116.2	1.2	5529	2	PF47010	AF007010 Plasmodiu
21	116.2	1.2	5529	6	AX428965	AX428965 Sequence
22	116.2	1.2	5529	6	AX536317	AX536317 Sequence
23	116.2	1.2	6152	6	AS8932	AS8932 Sequence 1
24	116.2	1.2	6152	6	AR131271	AR131271 Sequence
25	113.8	1.2	192275	9	AC164627	AC164627 Mus muscu
26	113.8	1.2	203395	14	AC164629	AC164629 Mus muscu
27	113.8	1.2	203787	14	AC153503	AC153503 Mus muscu
28	107.6	1.1	187543	14	AC141261	AC141261 Homo sapi
29	107.4	1.1	208332	9	AC122227	AC122227 Mus muscu
30	106.2	1.1	216959	9	AC116998	AC116998 Mus muscu
31	105.6	1.1	16377	2	AF202180	AF202180 Plasmodiu
32	105.2	1.1	110000	14	AC129389_1	Continuation (2 of
33	104.2	1.1	2069	6	E10125	E10125 DNA encodin
34	104.2	1.1	3399	6	E10126	E10126 DNA encodin
35	102.2	1.1	179685	9	AC122056	AC122056 Mus muscu
36	101.6	1.1	233713	14	AC133034	AC133034 Rattus no
37	101.4	1.0	228713	5	BX890609	BX890609 Zebrafish
38	100.4	1.0	4102	2	AF270648	AF270648 Plasmodiu
39	98.6	1.0	1701	6	CS075468	CS075468 Sequence
40	98.6	1.0	253001	6	CS082924	CS082924 Sequence
41	98.6	1.0	253001	2	AE014834	AE014834 Plasmodiu
42	97.4	1.0	140289	5	BX005458	BX005458 Zebrafish
43	97.4	1.0	237549	14	AC112451	AC112451 Rattus no
44	97	1.0	1964	2	AY2750284	AY275028 Plasmodiu
45	94.6	1.0	239813	14	AC098356	AC098356 Rattus no

## ALIGNMENTS

RESULT 1	AR442020	9726 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR442020	Sequence 8 from patent US 6669940.			
DEFINITION	AR442020				
ACCESSION	AR442020				
VERSION	AR442020.1	GI:42669171			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 9726)				
AUTHORS	Nagaraja,T.G., Stewart,G.C., Narayanan,S.K. and Chengappa,M.M.				
TITLE	Recombinant fusobacterium necrophorum leukotoxin vaccine and preparation thereof				
JOURNAL	Patent: US 6669940-A 8 30-DEC-2003;				
FEATURES	Kansas University Research Foundation; Manhattan, KS				
source	Location/Qualifiers				
	1..9726				
ORIGIN	/organism="Unknown"				
	/mol_type="genomic DNA"				
Query Match	100.0%;	Score 9726;	DB 6;	Length 9726;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 9726;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ATGAGCGGATCAAAATTAAGTTCAGAGGACAGGAGAGGATTCAGATTCTTAAAAA	60		
DB	1	ATGAGCGGATCAAAATTAAGTTCAGAGGACAGGAGAGGATTCAGATTCTTAAAAA	60		
QY	61	GTGTTAATGATTTGGGATTTGTTGATTAACATATGACGTTAGGGCTTAATGATCAATC	120		
DB	61	GTGTTAATGATTTGGGATTTGTTGATTAACATATGACGTTAGGGCTTAATGATCAATC	120		
QY	121	ACCGGACGACGAAATTTTGGAACTAAATTAAGGATTAATGATTCATTAATCT	180		
DB	121	ACCGGACGACGAAATTTTGGAACTAAATTAAGGATTAATGATTCATTAATCT	180		
QY	181	ACAAACAGATTCAGAGGAGGAGGAGGAGGATTCAGTTTAAAGATTTCCTTAAACAGAA	240		
DB	181	ACAAACAGATTCAGAGGAGGAGGAGGAGGATTCAGTTTAAAGATTTCCTTAAACAGAA	240		

QY 241 AATAATATAGCAAAATCTATATTTTGGGAAAAAAGATAGTACGGGGTAATAATCTTTT 300  
 Db 241 AATAATATAGCAAAATCTATATTTTGGGAAAAAAGATAGTACGGGGTAATAATCTTTT 300  
 QY 301 AACTTTGTCAATGGAATAATTTGAAGTGAATGGATTTATCAACGGAATTTGAGAAAAATAA 360  
 Db 301 AACTTTGTCAATGGAATAATTTGAAGTGAATGGATTTATCAACGGAATTTGAGAAAAATAA 360  
 QY 361 ATTGAGAGAAATTTATATTTCTTAAGCTCGAAGGGATGGCAGTAGAGAAAAATGGAAT 420  
 Db 361 ATTGAGAGAAATTTATATTTCTTAAGCTCGAAGGGATGGCAGTAGAGAAAAATGGAAT 420  
 QY 421 ATCAATGCTGTTCTTTTCATTTCTATATTTCCAAAACAAGATGATTTTAAAGAGCTTGG 480  
 Db 421 ATCAATGCTGTTCTTTTCATTTCTATATTTCCAAAACAAGATGATTTTAAAGAGCTTGG 480  
 QY 481 GAAAGAACCAACATGTGTAAAGTTTAAATGCAATCTCCAGTAGATGAAAAAGTAA 540  
 Db 481 GAAAGAACCAACATGTGTAAAGTTTAAATGCAATCTCCAGTAGATGAAAAAGTAA 540  
 QY 541 ATTCCATTGAATCCGAATGGAAGCATTTACGTTGAAAGAAAAATCAATGCTGTGAAGGC 600  
 Db 541 ATTCCATTGAATCCGAATGGAAGCATTTACGTTGAAAGAAAAATCAATGCTGTGAAGGC 600  
 QY 601 ATCGGTTTATATGCGCGGATATTTAGATTGAAGATGCAATCTAATAAGACAGAAAT 660  
 Db 601 ATCGGTTTATATGCGCGGATATTTAGATTGAAGATGCAATCTAATAAGACAGAAAT 660  
 QY 661 ACAGATTTTAAATTTAGTCAATATTTAGTCAATTAATTTCTGTGACCGAGAT 720  
 Db 661 ACAGATTTTAAATTTAGTCAATATTTAGTCAATTAATTTCTGTGACCGAGAT 720  
 QY 721 TTTAAAGCTACCAAGACAAATCTGAGATATTTCTTCACTCAGTCACTAGATTTCTCT 780  
 Db 721 TTTAAAGCTACCAAGACAAATCTGAGATATTTCTTCACTCAGTCACTAGATTTCTCT 780  
 QY 781 CAAAAAGCTATGGAAAAAATTCATCTGTGAAAGAGATAGAGATATGTAAAGGA 840  
 Db 781 CAAAAAGCTATGGAAAAAATTCATCTGTGAAAGAGATAGAGATATGTAAAGGA 840  
 QY 841 AATACCAAGCAAAATTTTGAATGTGATGCTGTATTTGAAAGCAGATGAAATTTAAAAAT 900  
 Db 841 AATACCAAGCAAAATTTTGAATGTGATGCTGTATTTGAAAGCAGATGAAATTTAAAAAT 900  
 QY 901 AGTGGAAAGCTACAAATGGAGATTATTAAGAAAGAGGGAAAAAGAACTTATAC 960  
 Db 901 AGTGGAAAGCTACAAATGGAGATTATTAAGAAAGAGGGAAAAAGAACTTATAC 960  
 QY 961 ACTCCTTTAAGTTTATCAGATGTGAAGCTCCGTAAAGATTAATTAAGAAAAAGTCATA 1020  
 Db 961 ACTCCTTTAAGTTTATCAGATGTGAAGCTCCGTAAAGATTAATTAAGAAAAAGTCATA 1020  
 QY 1021 GGAAGAAATGTTGACATTAACGTGAGAAAGAAATTTCTATGATGCACTTAACTTACT 1080  
 Db 1021 GGAAGAAATGTTGACATTAACGTGAGAAAGAAATTTCTATGATGCACTTAACTTACT 1080  
 QY 1081 AAGCTTGCAAGAGACTTTTAGCTTTGTATCAGGTTCTATTTCTCTATCAATTTAAAT 1140  
 Db 1081 AAGCTTGCAAGAGACTTTTAGCTTTGTATCAGGTTCTATTTCTCTATCAATTTAAAT 1140  
 QY 1141 GGAATTTTATGATTTATGCAAGTAAAGTCAAGTGTGCTTATTTGAAAGAAATGCCAAAGTC 1200  
 Db 1141 GGAATTTTATGATTTATGCAAGTAAAGTCAAGTGTGCTTATTTGAAAGAAATGCCAAAGTC 1200  
 QY 1201 GAAAGCAACGAAGGAAAGCAAAATATCTTCTTACAGTGAAGTAAAGCAACTATGGA 1260  
 Db 1201 GAAAGCAACGAAGGAAAGCAAAATATCTTCTTACAGTGAAGTAAAGCAACTATGGA 1260  
 QY 1261 GGAAGCAACGAAGGAAAGCAAAATATCTTCTTACAGTGAAGTAAAGCAACTATGGA 1320  
 Db 1261 GGAAGCAACGAAGGAAAGCAAAATATCTTCTTACAGTGAAGTAAAGCAACTATGGA 1320

QY 1321 CTCAGTATCGAGCGGGATATATTTCTGCAAAAAAGTAATTCOAATGTAATTAATGAAGA 1380  
 Db 1321 CTCAGTATCGAGCGGGATATATTTCTGCAAAAAAGTAATTCOAATGTAATTAATGAAGA 1380  
 QY 1381 GAAGTAAATTCGAAGGGAGAGAGATTTTCTTCAAAATCTGAAAAATCTATGATGCT 1440  
 Db 1381 GAAGTAAATTCGAAGGGAGAGAGATTTTCTTCAAAATCTGAAAAATCTATGATGCT 1440  
 QY 1441 TCTGTTCTGTTGGAACGATGAGAGATTCOAATTAAGTACTTTTCAGTATTTGTAAGC 1500  
 Db 1441 TCTGTTCTGTTGGAACGATGAGAGATTCOAATTAAGTACTTTTCAGTATTTGTAAGC 1500  
 QY 1501 GAAGGAAAAATTAATCTTCGTCAGATTGCTAAAGGACAAAGTAAATCGAAAAAG 1560  
 Db 1501 GAAGGAAAAATTAATCTTCGTCAGATTGCTAAAGGACAAAGTAAATCGAAAAAG 1560  
 QY 1561 GATGATGTAATTCGGAAGAGTGAAGCGATTAATTCOAATTCGAGCTGCTGTAAGGTGA 1620  
 Db 1561 GATGATGTAATTCGGAAGAGTGAAGCGATTAATTCOAATTCGAGCTGCTGTAAGGTGA 1620  
 QY 1621 TTGGGGGATAGTGGTAATGGGGTTGTGGCTGCAAAATATTTCTAATATATGCTTCTCC 1680  
 Db 1621 TTGGGGGATAGTGGTAATGGGGTTGTGGCTGCAAAATATTTCTAATATATGCTTCTCC 1680  
 QY 1681 CGTATGATGTAATGATGATGATCTACATGCAAGAAAGCACTAAATGTGAGAGCTCATAC 1740  
 Db 1681 CGTATGATGTAATGATGATGATCTACATGCAAGAAAGCACTAAATGTGAGAGCTCATAC 1740  
 QY 1741 ATTACTAATAATTAAGTTTCTGCCAAACAGATCTGATTTGGGAACCTTCCAGTTATGAT 1800  
 Db 1741 ATTACTAATAATTAAGTTTCTGCCAAACAGATCTGATTTGGGAACCTTCCAGTTATGAT 1800  
 QY 1801 GATCAGTTTATGAAATCAGTCAATCTAATAATCAATTTTATAGTCAATTAATGAGCTTT 1860  
 Db 1801 GATCAGTTTATGAAATCAGTCAATCTAATAATCAATTTTATAGTCAATTAATGAGCTTT 1860  
 QY 1861 GGAGGAGACAGTGTCAATGAGGAAATTAAGAAATGCTAACGAATTAATTTAGTGTGCT 1920  
 Db 1861 GGAGGAGACAGTGTCAATGAGGAAATTAAGAAATGCTAACGAATTAATTTAGTGTGCT 1920  
 QY 1921 GTGTGTGCAACCTAGCAATTCATTAATTAATTTCTGTTGTGGCAATGAGAGAGTGA 1980  
 Db 1921 GTGTGTGCAACCTAGCAATTCATTAATTAATTTCTGTTGTGGCAATGAGAGAGTGA 1980  
 QY 1981 AGACTTTCTTCAAGAGTGAAGAGAGATGTAATGAGGCAATTAATGAGCTCAAAATCTT 2040  
 Db 1981 AGACTTTCTTCAAGAGTGAAGAGAGATGTAATGAGGCAATTAATGAGCTCAAAATCTT 2040  
 QY 2041 CGAGGACTAGCTCAAGTGAAGTGTGCTGTACGAAGGAAAGAAAAAGAACTTAT 2100  
 Db 2041 CGAGGACTAGCTCAAGTGAAGTGTGCTGTACGAAGGAAAGAAAAAGAACTTAT 2100  
 QY 2101 GGAATGCAAGCTTTTATATGAAAACTATTAATAATTCGTTCTGTGCAATTTGCCAT 2160  
 Db 2101 GGAATGCAAGCTTTTATATGAAAACTATTAATAATTCGTTCTGTGCAATTTGCCAT 2160  
 QY 2161 CATGCTGAATTTGTGGAAGGAAAAATGTATCAACAGTGAATTAATTAATGATAT 2220  
 Db 2161 CATGCTGAATTTGTGGAAGGAAAAATGTATCAACAGTGAATTAATTAATGATAT 2220  
 QY 2221 AAAAAATCTTCAAAAAATGCAAAAGTCTGTATTTGTAATTAATGAACCTTTTAAAGAGCT 2280  
 Db 2221 AAAAAATCTTCAAAAAATGCAAAAGTCTGTATTTGTAATTAATGAACCTTTTAAAGAGCT 2280  
 QY 2281 TTTGAAAAAGAAAGCAAAATCTCCAGATATGATCCGAAGAGATTAATGAATCTATGAAAA 2340  
 Db 2281 TTTGAAAAAGAAAGCAAAATCTCCAGATATGATCCGAAGAGATTAATGAATCTATGAAAA 2340  
 QY 2341 TTAATGAATGATTTTCAAAAAATTTGATGAAAAAGCGAGCTTTTCTAATAGGTGA 2400  
 Db 2341 TTAATGAATGATTTTCAAAAAATTTGATGAAAAAGCGAGCTTTTCTAATAGGTGA 2400  
 QY 2401 AGAATGACAATTAATCTTCCGAGTGAACCTTCAAAAAAGGAACTGTATAGAAATGCA 2460

Db 2401 AGAATGACAAATTATCTTCCGGATGGAACCTTCAAAAAGGAACTGCTATAGAAATTGCA 2460  
Qy 2461 AACTATGTTGAGGGGAAATGAAAAATTAGGAAAAATTCCGAAAGATTTTAAAGCT 2520  
Db 2461 AACTATGTTGAGGGGAAATGAAAAATTAGGAAAAATTCCGAAAGATTTTAAAGCT 2520  
Qy 2521 TTTTCAGAGAGTATAGTGAAGTGAATTAAAGAACTTTGAATTTTACAGAGATAGGAAAT 2580  
Db 2521 TTTTCAGAGAGTATAGTGAAGTGAATTAAAGAACTTTGAATTTTACAGAGATAGGAAAT 2580  
Qy 2581 TATGCAAAATTTTCAACATTTTACCTTCCGAGCTAATGAGAAAGAGATTTTCTTCT 2640  
Db 2581 TATGCAAAATTTTCAACATTTTACCTTCCGAGCTAATGAGAAAGAGATTTTCTTCT 2640  
Qy 2641 GTGGAGAGAGCTGTTTCTGCGGTAGAACAGAGAAATTATGCAAGGTATCCGTTGGAAA 2700  
Db 2641 GTGGAGAGAGCTGTTTCTGCGGTAGAACAGAGAAATTATGCAAGGTATCCGTTGGAAA 2700  
Qy 2701 GGAGCTAAACTGCTGCAAAAAAAGATTAAATATTAAGCTATCAATTAAGCAGAAACA 2760  
Db 2701 GGAGCTAAACTGCTGCAAAAAAAGATTAAATATTAAGCTATCAATTAAGCAGAAACA 2760  
Qy 2761 GTGAAATTTAGTTGAAATATTTGAACTTGGAGAGAGAGTACATCCGAAAGTCAGTCGA 2820  
Db 2761 GTGAAATTTAGTTGAAATATTTGAACTTGGAGAGAGTACATCCGAAAGTCAGTCGA 2820  
Qy 2821 GGAAGATTAATGTTTCAAAAGATCGAAAAATTCAGCTATGCTAGAGCTTAAAGAAAAGCT 2880  
Db 2821 GGAAGATTAATGTTTCAAAAGATCGAAAAATTCAGCTATGCTAGAGCTTAAAGAAAAGCT 2880  
Qy 2881 GAATTTATCGAGAAAAATTTAATGAGATGCAATTCGAACAGACTTTTTCATGAGCGGA 2940  
Db 2881 GAATTTATCGAGAAAAATTTAATGAGATGCAATTCGAACAGACTTTTTCATGAGCGGA 2940  
Qy 2941 TCTTTTAAATGCTGAGTGCAGGTGGAATGCAATCAATGGAAGGAAATATAGTGAAGCT 3000  
Db 2941 TCTTTTAAATGCTGAGTGCAGGTGGAATGCAATCAATGGAAGGAAATATAGTGAAGCT 3000  
Qy 3001 ATCAGTAAAGCAAGAGTTTCCATTTGATGACGAAGCATATTTGAAACCTATTAATAAAAT 3060  
Db 3001 ATCAGTAAAGCAAGAGTTTCCATTTGATGACGAAGCATATTTGAAACCTATTAATAAAAT 3060  
Qy 3061 GCTTTTAAACAGTAAAGATGATATCTTCTGTTTGAATGCTGCCGTTCAACGGGAATCGGA 3120  
Db 3061 GCTTTTAAACAGTAAAGATGATATCTTCTGTTTGAATGCTGCCGTTCAACGGGAATCGGA 3120  
Qy 3121 ACGAAAAATGCGGCGGTGCGGGGTTGCTGTTGCGGTAAATGATATATTTCAAAACAA 3180  
Db 3121 ACGAAAAATGCGGCGGTGCGGGGTTGCTGTTGCGGTAAATGATATATTTCAAAACAA 3180  
Qy 3181 GCTTCCATTTGAAGATATATGACGAAGCAAAAGTAAATATGATTAAGATTAAGATGAA 3240  
Db 3181 GCTTCCATTTGAAGATATATGACGAAGCAAAAGTAAATATGATTAAGATTAAGATGAA 3240  
Qy 3241 GTTACGTAATCTGCGGAATCTTTTGAAGTATGATGCAAAAAAGCCGGAACAATCAACGT 3300  
Db 3241 GTTACGTAATCTGCGGAATCTTTTGAAGTATGATGCAAAAAAGCCGGAACAATCAACGT 3300  
Qy 3241 GTTACGTAATCTGCGGAATCTTTTGAAGTATGATGCAAAAAAGCCGGAACAATCAACGT 3300  
Db 3241 GTTACGTAATCTGCGGAATCTTTTGAAGTATGATGCAAAAAAGCCGGAACAATCAACGT 3300  
Qy 3301 ATTTCTGTTGCGGAGGAATTAATAAGTTGAAGTAAACGAGTGAAGAAAAACGAAA 3360  
Db 3301 ATTTCTGTTGCGGAGGAATTAATAAGTTGAAGTAAACGAGTGAAGAAAAACGAAA 3360  
Qy 3361 TCAGAGAAAGAACAGAGGATTTTGTGCAAAATCGGAAACAAAGTGAAGCTGTGTAA 3420  
Db 3361 TCAGAGAAAGAACAGAGGATTTTGTGCAAAATCGGAAACAAAGTGAAGCTGTGTAA 3420  
Qy 3421 AATTAATTTGAGATATGATATCTTAAACGAAAAAATTAACAATTAATTTCTGAA 3480  
Db 3421 AATTAATTTGAGATATGATATCTTAAACGAAAAAATTAACAATTAATTTCTGAA 3480  
Qy 3481 GAGTAAAAAAGCGGGGAATCTTCTCGAAGCTTCTCATCTCCGATTAAGGACG 3540  
Db 3481 GAGTAAAAAAGCGGGGAATCTTCTCGAAGCTTCTCATCTCCGATTAAGGACG 3540

Db 3481 GAGTAAAAAAGCGGGGAATCTTCTCGAAGCTTCTCATCTCCGATTAAGGACG 3540  
Qy 3541 TCTTTCACTTTGGAGAGCTTCTGAAAGTGTCTTCAATTAATTAATAAGAAACATCT 3600  
Db 3541 TCTTTCACTTTGGAGAGCTTCTGAAAGTGTCTTCAATTAATTAATAAGAAACATCT 3600  
Qy 3601 GCTGTCTGATGAGTAAAGATTAATTTGAAGGAGCAATTAATAAGGTAGAGTACT 3660  
Db 3601 GCTGTCTGATGAGTAAAGATTAATTTGAAGGAGCAATTAATAAGGTAGAGTACT 3660  
Qy 3661 TCTTCTGATTTACTTTTGTGAGACATGAGGCGGATCTGCTGCACTTCACTGAAATCAT 3720  
Db 3661 TCTTCTGATTTACTTTTGTGAGACATGAGGCGGATCTGCTGCACTTCACTGAAATCAT 3720  
Qy 3721 ATTGAAGTGAAGAAATAGCAACATCAGTCTGTTAGCTGAGCGGCTGCTTAATAT 3780  
Db 3721 ATTGAAGTGAAGAAATAGCAACATCAGTCTGTTAGCTGAGCGGCTGCTTAATAT 3780  
Qy 3781 ATTCAAGTAAACCAAGTCTTGTGTTAAAAATAGTATATTGAAATGCAATTAATT 3840  
Db 3781 ATTCAAGTAAACCAAGTCTTGTGTTAAAAATAGTATATTGAAATGCAATTAATT 3840  
Qy 3841 AAAGTAAATGCTTTGATGAGAGAACTCAAGTACAGAGAGAGAGCTTTGAGACAGT 3900  
Db 3841 AAAGTAAATGCTTTGATGAGAGAACTCAAGTACAGAGAGAGAGCTTTGAGACAGT 3900  
Qy 3901 AAAGTAAATGCTTTGATGAGAGAACTCAAGTACAGAGAGAGAGCTTTGAGACAGT 3960  
Db 3901 AAAGTAAATGCTTTGATGAGAGAACTCAAGTACAGAGAGAGAGCTTTGAGACAGT 3960  
Qy 3961 GTGAACTAATGAGTTTCTGCAAAATCAGAAAAATTAACAGTACAGAGAGAACTGAAAC 4020  
Db 3961 GTGAACTAATGAGTTTCTGCAAAATCAGAAAAATTAACAGTACAGAGAGAACTGAAAC 4020  
Qy 4021 CAAAAATGAGTGTGATGATCACTGCTTATCAAGCGGACACCAAGTGAAGAGCTTTA 4080  
Db 4021 CAAAAATGAGTGTGATGATCACTGCTTATCAAGCGGACACCAAGTGAAGAGCTTTA 4080  
Qy 4081 AATTTTAAAGCTGGAAGAGTCAATGGAATGGAAGAGAGAGAGAGAGAGAGAGAGAG 4140  
Db 4081 AATTTTAAAGCTGGAAGAGTCAATGGAATGGAAGAGAGAGAGAGAGAGAGAGAGAG 4140  
Qy 4141 AACCAAAAGTAAATGCTTCTATTTAGTGGAGATATATCAAGCTTAATGAGAGAGAG 4200  
Db 4141 AACCAAAAGTAAATGCTTCTATTTAGTGGAGATATATCAAGCTTAATGAGAGAGAG 4200  
Qy 4201 GCAAAAGCTCTTTTAGCAACCACTCAAGTGAAGTGAAGAGAGAGAGAGAGAGAG 4260  
Db 4201 GCAAAAGCTCTTTTAGCAACCACTCAAGTGAAGTGAAGAGAGAGAGAGAGAGAG 4260  
Qy 4261 AGTTCTGAGCGGAGTTAGGAATTAATCAAGGGGCTTTCTGTCAATTAAGATGCAAT 4320  
Db 4261 AGTTCTGAGCGGAGTTAGGAATTAATCAAGGGGCTTTCTGTCAATTAAGATGCAAT 4320  
Qy 4321 GAGGTGAAGCTAGCGTTGATTAATCTTCACTGAAGAGAGTAAATGCAATGTCAT 4380  
Db 4321 GAGGTGAAGCTAGCGTTGATTAATCTTCACTGAAGAGAGTAAATGCAATGTCAT 4380  
Qy 4381 GCCAAAGATGTCAAAAGAGTTCTGATCTAGCAAAAGAAATATCAGGCTTAAATGGA 4440  
Db 4381 GCCAAAGATGTCAAAAGAGTTCTGATCTAGCAAAAGAAATATCAGGCTTAAATGGA 4440  
Qy 4441 AAAAGTAAATTAATTTAAGATGCTGTATTAATACAGCTGGAATGCTTAATAG 4500  
Db 4441 AAAAGTAAATTAATTTAAGATGCTGTATTAATACAGCTGGAATGCTTAATAG 4500  
Qy 4501 AAGGAACAATGAGAAAAAGCAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560  
Db 4501 AAGGAACAATGAGAAAAAGCAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560  
Qy 4561 TCGGTTGCTGGAACGAGTAAATCCGCTGAGAGAGTATGCTATGCAATATCTGTTAA 4620  
Db 4561 TCGGTTGCTGGAACGAGTAAATCCGCTGAGAGAGTATGCTATGCAATATCTGTTAA 4620

QY 4621 AATTAATTTAAAGCAATTTGAGTGAAGCAATTAAGAAAGCGAGAGATAAATTCAT 4680  
| | | | |  
Db 4621 AATTAATTTAAAGCAAAATTGAGTGAAGCAATTAAGAAAGCGAGAGATAAATTCAT 4680  
| | | | |  
QY 4681 GCGAAACATGTAAATGTGAAGCGCAAAATCATCTGTTGTTGTGAATCGGCTTCTGGA 4740  
| | | | |  
Db 4681 GCGAAACATGTAAATGTGAAGCGCAAAATCATCTGTTGTTGTGAATCGGCTTCTGGA 4740  
| | | | |  
QY 4741 CTTCGATCAGCAAGAATGCTTTTCAGGAATGGGATCTGGAGCANTGGGAAGACTTATCA 4800  
| | | | |  
Db 4741 CTTCGATCAGCAAGAATGCTTTTCAGGAATGGGATCTGGAGCANTGGGAAGACTTATCA 4800  
| | | | |  
QY 4801 AATGACACGATTTGCAAAAGGTGATAAAGGAAGAAATTTCTGATGATTCCTTAAATGTAAC 4860  
| | | | |  
Db 4801 AATGACACGATTTGCAAAAGGTGATAAAGGAAGAAATTTCTGATGATTCCTTAAATGTAAC 4860  
| | | | |  
QY 4861 GCAAAATTAATTCATCTTGGGGTGAATGTTGCGGGAACCATTTGCCGGTTCCTTCTTACG 4920  
| | | | |  
Db 4861 GCAAAATTAATTCATCTTGGGGTGAATGTTGCGGGAACCATTTGCCGGTTCCTTCTTACG 4920  
| | | | |  
QY 4921 GCGGTAAGAGCTGCTTTTGGCAATATCTCTTCATTAATTAACCTGCTGCTTGAATTA 4980  
| | | | |  
Db 4921 GCGGTAAGAGCTGCTTTTGGCAATATCTCTTCATTAATTAACCTGCTGCTTGAATTA 4980  
| | | | |  
QY 4981 GGAACGAAAGTAAATCCTTTTAGTGAAGAAATACAAAGATCATGTAAGCTTTGAAT 5040  
| | | | |  
Db 4981 GGAACGAAAGTAAATCCTTTTAGTGAAGAAATACAAAGATCATGTAAGCTTTGAAT 5040  
| | | | |  
QY 5041 GATTCATATTTACAAAGCTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGCTGGAATC 5100  
| | | | |  
Db 5041 GATTCATATTTACAAAGCTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGCTGGAATC 5100  
| | | | |  
QY 5101 GGAGGAATGTATCTGTCAATGCTGCTGATGAAAGGAAAGCTTTAGTATGATTC 5160  
| | | | |  
Db 5101 GGAGGAATGTATCTGTCAATGCTGCTGATGAAAGGAAAGCTTTAGTATGATTC 5160  
| | | | |  
QY 5161 GAGTTGAAGAGATTAATCTTTCAATGTAGATGCAAAAGATCAAAACAAATTAATTA 5220  
| | | | |  
Db 5161 GAGTTGAAGAGATTAATCTTTCAATGTAGATGCAAAAGATCAAAACAAATTAATTA 5220  
| | | | |  
QY 5221 ATTGCCGGAATGCAATGGAAGAAAGCGCTGCTGAGTGGAGCAACAGTTCTCATACA 5280  
| | | | |  
Db 5221 ATTGCCGGAATGCAATGGAAGAAAGCGCTGCTGAGTGGAGCAACAGTTCTCATACA 5280  
| | | | |  
QY 5281 AATATTTGAAACAAATCAGTTATAGCTATTTGTAAABAAAGTAAATTAACAAGCGAAT 5340  
| | | | |  
Db 5281 AATATTTGAAACAAATCAGTTATAGCTATTTGTAAABAAAGTAAATTAACAAGCGAAT 5340  
| | | | |  
QY 5341 GATCAAGATAGAAATATCAATGTGACTGCAAAAGATTAATCAATCAATCAATCAAT 5400  
| | | | |  
Db 5341 GATCAAGATAGAAATATCAATGTGACTGCAAAAGATTAATCAATCAATCAATCAAT 5400  
| | | | |  
QY 5401 GCAAGTCGAGTGGAGAGCAAAAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTACC 5460  
| | | | |  
Db 5401 GCAAGTCGAGTGGAGAGCAAAAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTACC 5460  
| | | | |  
QY 5461 TTGAATTAAGACAGTTTCTCTCATGTTGATCAAACTGATTAAGCAAAATTAAGGA 5520  
| | | | |  
Db 5461 TTGAATTAAGACAGTTTCTCTCATGTTGATCAAACTGATTAAGCAAAATTAAGGA 5520  
| | | | |  
QY 5521 GAAATTAATGAAATTAAGAAAGCAATGTTAATGTTCTAGCTGAAATACGAGTCA 5580  
| | | | |  
Db 5521 GAAATTAATGAAATTAAGAAAGCAATGTTAATGTTCTAGCTGAAATACGAGTCA 5580  
| | | | |  
QY 5581 GTGTGTCACAAATGCGACAGTCTTTCGAGGCAAGTGAACAAGCTGAGTAGAGAGTGA 5640  
| | | | |  
Db 5581 GTGTGTCACAAATGCGACAGTCTTTCGAGGCAAGTGAACAAGCTGAGTAGAGAGTGA 5640  
| | | | |  
QY 5641 GTAGCAGTTAATTAATTAATCAAAAAATCTTCTGCAATATAAAAAATGTAAGTCA 5700  
| | | | |  
Db 5641 GTAGCAGTTAATTAATTAATCAAAAAATCTTCTGCAATATAAAAAATGTAAGTCA 5700  
| | | | |

QY 5701 GTACGAATAGCTTTGTGTAAAAAGCAAAATCTCATCTATTAATAACATTTGAAATGGA 5760  
| | | | |  
Db 5701 GTACGAATAGCTTTGTGTAAAAAGCAAAATCTCATCTATTAATAACATTTGAAATGGA 5760  
| | | | |  
QY 5761 GCTGAGTTGAGCTGAGAGAGCTGAGAGTGAAGTTCGTATGACAGTGAATTAAGATTGTA 5820  
| | | | |  
Db 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGAGTGAAGTTCGTATGACAGTGAATTAAGATTGTA 5820  
| | | | |  
QY 5821 AATTAATGAGTATGAGAAATTAATCATGCAAAAATCACTGCGAAGGAAATGTCGAGAT 5880  
| | | | |  
Db 5821 AATTAATGAGTATGAGAAATTAATCATGCAAAAATCACTGCGAAGGAAATGTCGAGAT 5880  
| | | | |  
QY 5881 ATTACAGAGTCTGATGCGGTATTTGCTAATTAATGAGGAACAGTGTCTGAGATGCCCC 5940  
| | | | |  
Db 5881 ATTACAGAGTCTGATGCGGTATTTGCTAATTAATGAGGAACAGTGTCTGAGATGCCCC 5940  
| | | | |  
QY 5941 GCAGCAATAGAGGCTCAACCGAGTGAATGAATTAACAGGATCTACAAAGCATATGTA 6000  
| | | | |  
Db 5941 GCAGCAATAGAGGCTCAACCGAGTGAATGAATTAACAGGATCTACAAAGCATATGTA 6000  
| | | | |  
QY 6001 AAAGATTTACAGATTTGCTTAAAGAAAGAGATGATTAATTAATTAATTAATTAATTA 6060  
| | | | |  
Db 6001 AAAGATTTACAGATTTGCTTAAAGAAAGAGATGATTAATTAATTAATTAATTAATTA 6060  
| | | | |  
QY 6061 GTAGATTAAGTGTAGATTAAGTATTTCAAAATCTTAATTAATTAACGAAGCTTATCACA 6120  
| | | | |  
Db 6061 GTAGATTAAGTGTAGATTAAGTATTTCAAAATCTTAATTAATTAACGAAGCTTATCACA 6120  
| | | | |  
QY 6121 AAAAGAAAAATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6180  
| | | | |  
Db 6121 AAAAGAAAAATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6180  
| | | | |  
QY 6181 AAATCTTTATTTGGAAGAAATGCGCTGCTTCAAGACAAAGCGAGGAGGAGAACTGTAT 6240  
| | | | |  
Db 6181 AAATCTTTATTTGGAAGAAATGCGCTGCTTCAAGACAAAGCGAGGAGGAGAACTGTAT 6240  
| | | | |  
QY 6241 ATCAACAAAGTTTATGAGAAACAGAACTCTTTAGTAAATTTCTAATTAATTAATTA 6300  
| | | | |  
Db 6241 ATCAACAAAGTTTATGAGAAACAGAACTCTTTAGTAAATTTCTAATTAATTAATTA 6300  
| | | | |  
QY 6301 CATTAATTTCTGTAATTAACAGAGATTAACAGAAATTCATTCGAGATGTAAGTCTGTTG 6360  
| | | | |  
Db 6301 CATTAATTTCTGTAATTAACAGAGATTAACAGAAATTCATTCGAGATGTAAGTCTGTTG 6360  
| | | | |  
QY 6361 GTTGTGGAATTTGTAAGAGTGGAGCTTCTCTGATACCAATTAATTAATTAATTAAT 6420  
| | | | |  
Db 6361 GTTGTGGAATTTGTAAGAGTGGAGCTTCTCTGATACCAATTAATTAATTAATTAAT 6420  
| | | | |  
QY 6421 AAGACAAGATTTGAAAAAATTAACAATGTCGTATGAAGTTTCGAGAAAGAGCTGAAAT 6480  
| | | | |  
Db 6421 AAGACAAGATTTGAAAAAATTAACAATGTCGTATGAAGTTTCGAGAAAGAGCTGAAAT 6480  
| | | | |  
QY 6481 ACAGCAGATTTCTAGCAAGAAATTTCTCTTTTGAAGTTCGAGATGCGACGCGGGTA 6540  
| | | | |  
Db 6481 ACAGCAGATTTCTAGCAAGAAATTTCTCTTTTGAAGTTCGAGATGCGACGCGGGTA 6540  
| | | | |  
QY 6541 GGAAGCCGAGTGGAGAGAAACGTTTCGCTAAATCAATTTGCAAGGAAGACGGAATGAT 6600  
| | | | |  
Db 6541 GGAAGCCGAGTGGAGAGAAACGTTTCGCTAAATCAATTTGCAAGGAAGACGGAATGAT 6600  
| | | | |  
QY 6601 GTGGAAGAAGCAAGATTTTGTATTAATAAAGCTGAGATTAACAGCAAAAGCTTATAGTCT 6660  
| | | | |  
Db 6601 GTGGAAGAAGCAAGATTTTGTATTAATAAAGCTGAGATTAACAGCAAAAGCTTATAGTCT 6660  
| | | | |  
QY 6661 GTTGCAATTTGAAATGCGCAGTCCGAGTGGCTGCAAAAGAGCTGGAATTTGAGCAGCA 6720  
| | | | |  
Db 6661 GTTGCAATTTGAAATGCGCAGTCCGAGTGGCTGCAAAAGAGCTGGAATTTGAGCAGCA 6720  
| | | | |  
QY 6721 GTGCGAGTTACCAAGATGAATCAACAGAGAGCAAGAGGAAATTTCTTAAATTTATG 6780  
| | | | |  
Db 6721 GTGCGAGTTACCAAGATGAATCAACAGAGAGCAAGAGGAAATTTCTTAAATTTATG 6780  
| | | | |  
QY 6781 ACTCGAAACAAGTTAGATGTATAGCAAAATGAGATTAATCAAGTACTGGAATCGAT 6840  
| | | | |

Db 6781 ACTCGAAACAAGTTAGATCTATACGAGAAATGAGTAATAATCAGATCTCGAATCCGT 6840  
Oy 6841 TCAGCCGAGAGCTGGAATCTTTCAGCCGAGATCTGAGTGTCTTCTGTAATATAT 6900  
Db 6841 TCAGCCGAGAGCTGGAATCTTTCAGCCGAGATCTGAGTGTCTTCTGTAATATAT 6900  
Oy 6901 GCMAATTAAGTGAAGAACAGATATCGATCTTACTCTTCTTCTATCTATGTAAT 6960  
Db 6901 GCMAATTAAGTGAAGAACAGATATCGATCTTACTCTTCTTCTATCTATGTAAT 6960  
Oy 6961 GTPAAAGCTCTTAAATTAATTTTTCGAATCTTTCAGCCGCTGAGAGAGCCGAGCTT 7020  
Db 6961 GTPAAAGCTCTTAAATTAATTTTTCGAATCTTTCAGCCGCTGAGAGAGCCGAGCTT 7020  
Oy 7021 GCAGCAGTTACCGAGAGTGTCTTCTGTTAACTATTAATAGTTCTGTAGTCTGAGTT 7080  
Db 7021 GCAGCAGTTACCGAGAGTGTCTTCTGTTAACTATTAATAGTTCTGTAGTCTGAGTT 7080  
Oy 7081 CACAAATACCTGATTTGACTTCCGTACGAGAAAAAGTAAATGTAAACGCAAAAGAGAA 7140  
Db 7081 CACAAATACCTGATTTGACTTCCGTACGAGAAAAAGTAAATGTAAACGCAAAAGAGAA 7140  
Oy 7141 AAAAATATTAGCAAAACAGCAGCAAAATGAGGAATGAGAGAGCAGCAATCGAGCCAT 7200  
Db 7141 AAAAATATTAGCAAAACAGCAGCAAAATGAGGAATGAGAGAGCAGCAATCGAGCCAT 7200  
Oy 7201 GTCTTGTAATTAATTTTGGAAACGCTGTAGAGATAGAAAAATTTCTGAGAGAAAAAG 7260  
Db 7201 GTCTTGTAATTAATTTTGGAAACGCTGTAGAGATAGAAAAATTTCTGAGAGAAAAAG 7260  
Oy 7261 ACAGAAAGTTTAAAACTTTAGACGAAGTTAACAAAGAACAGATTAATAATGAT 7320  
Db 7261 ACAGAAAGTTTAAAACTTTAGACGAAGTTAACAAAGAACAGATTAATAATGAT 7320  
Oy 7321 GCTAGCAAAAAATCTTACATCAGAGGATTTCTACAGAGATACCTCTGTAAGAGG 7380  
Db 7321 GCTAGCAAAAAATCTTACATCAGAGGATTTCTACAGAGATACCTCTGTAAGAGG 7380  
Oy 7381 GATAGAGAGATACCTCAGGAGAGAAATTAAGCCATTTGTGAGACTTCTGATATAT 7440  
Db 7381 GATAGAGAGATACCTCAGGAGAGAAATTAAGCCATTTGTGAGACTTCTGATATAT 7440  
Oy 7441 GAAAAAATGTAGATTTTACACAGAGACAGAAATATATCTTCTATCTGCTGTTG 7500  
Db 7441 GAAAAAATGTAGATTTTACACAGAGACAGAAATATATCTTCTATCTGCTGTTG 7500  
Oy 7501 GGAACCTGCAAGTCTTCCGATCAGAGAACGTGCGATTAACAATATTTAAAGAAAT 7560  
Db 7501 GGAACCTGCAAGTCTTCCGATCAGAGAACGTGCGATTAACAATATTTAAAGAAAT 7560  
Oy 7561 TCCGAGTACTACTGTGAAAAATCTTTTGTGAAAGCAGCTGAAAAAGTAAATGTTAGATG 7620  
Db 7561 TCCGAGTACTACTGTGAAAAATCTTTTGTGAAAGCAGCTGAAAAAGTAAATGTTAGATG 7620  
Oy 7621 GATATTTACAGAAATGTGCTTTTACAGCATATCAAGTCTCTAGAGAGATTTGGAAAT 7680  
Db 7621 GATATTTACAGAAATGTGCTTTTACAGCATATCAAGTCTCTAGAGAGATTTGGAAAT 7680  
Oy 7681 GAGAGCTGCTATCAGAAATTAATTTCTAATGGAAGATCAATATCTGATTAATAATTC 7740  
Db 7681 GAGAGCTGCTATCAGAAATTAATTTCTAATGGAAGATCAATATCTGATTAATAATTC 7740  
Oy 7741 AAGCTATTAGAAAAATATATTGATGTTATTTAAAGATTAATCGAATTTGAGAGCGAA 7800  
Db 7741 AAGCTATTAGAAAAATATATTGATGTTATTTAAAGATTAATCGAATTTGAGAGCGAA 7800  
Oy 7801 GCAAAAGATTAACCGTAGAGGCTGCTGCGAGAGCCATTTCTCAAAAGCAAGAT 7860  
Db 7801 GCAAAAGATTAACCGTAGAGGCTGCTGCGAGAGCCATTTCTCAAAAGCAAGAT 7860  
Oy 7861 GAAATGAATTCAGAGGTGAAATTTGAGAGAGTATTTCAATGAGAAAAATAGAGTACT 7920

Db 7861 GAAATGAATTCAGAGGTGAAATTTGAGAGAGTATTTTCAATGAGAAAAATAGATACT 7920  
Oy 7921 AGCCCTTAAAGAAATTGAAAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTG 7980  
Db 7921 AGCCCTTAAAGAAATTGAAAGAAATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTG 7980  
Oy 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGAGGGGCGAGAAATTAATTTCCGA 8040  
Db 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGAGGGGCGAGAAATTAATTTCCGA 8040  
Oy 8041 GCAAAAGATGCCGAGAGCTTATTTGAAAGTATGACAAATCCGGAAGAAATTTTT 8100  
Db 8041 GCAAAAGATGCCGAGAGCTTATTTGAAAGTATGACAAATCCGGAAGAAATTTTT 8100  
Oy 8101 CATGCAATTAATGTGAATATGGAAGCAACATAAATGAAAGTAAACGAGCTTCTAA 8160  
Db 8101 CATGCAATTAATGTGAATATGGAAGCAACATAAATGAAAGTAAACGAGCTTCTAA 8160  
Oy 8161 GCAGTAAACAGTTCTGTATTGGAGAGAGTGGAGTCAACCAAGGCAAGACTACTGTCGA 8220  
Db 8161 GCAGTAAACAGTTCTGTATTGGAGAGAGTGGAGTCAACCAAGGCAAGACTACTGTCGA 8220  
Oy 8221 GGTAAACCTATGTAGAAAGTTGAGAGAGAAATTTGTCAGAACAAATCGATTGAATGCA 8280  
Db 8221 GGTAAACCTATGTAGAAAGTTGAGAGAGAAATTTGTCAGAACAAATCGATTGAATGCA 8280  
Oy 8281 ATTTTAAAGTGAAGGTTTGAATGAATTAAGTAACTGCTAAATCTTCTGATATCA 8340  
Db 8281 ATTTTAAAGTGAAGGTTTGAATGAATTAAGTAACTGCTAAATCTTCTGATATCA 8340  
Oy 8341 GGAATGAGAGAGAAATTCGAGAGAGAGTGAATCTTCTACAGACCAAAAGTAACT 8400  
Db 8341 GGAATGAGAGAGAAATTCGAGAGAGAGTGAATCTTCTACAGACCAAAAGTAACT 8400  
Oy 8401 GAATCCGTAGTTCGTTTACGAAAGCAAGATTAGAAATATGATTTACCAAAAAATAT 8460  
Db 8401 GAATCCGTAGTTCGTTTACGAAAGCAAGATTAGAAATATGATTTACCAAAAAATAT 8460  
Oy 8461 ATTTGAGAGTCAATGCTCTTCTTAAATGATTAACAAAGAAAGGCGAATTTAGAAATC 8520  
Db 8461 ATTTGAGAGTCAATGCTCTTCTTAAATGATTAACAAAGAAAGGCGAATTTAGAAATC 8520  
Oy 8521 TTAGCGGTAGCCGAGTGTGATGCAAGAGAACAAACAAAGCAATTTACAGATCAAAAG 8580  
Db 8521 TTAGCGGTAGCCGAGTGTGATGCAAGAGAACAAACAAAGCAATTTACAGATCAAAAG 8580  
Oy 8581 TTAACTTCTAACAATGTAAATGAGAGAAACGTATCTCAATCTGTCGAAAGCTTTGGCT 8640  
Db 8581 TTAACTTCTAACAATGTAAATGAGAGAAACGTATCTCAATCTGTCGAAAGCTTTGGCT 8640  
Oy 8641 AAAAATGAAATTTATGGAATATGTAAGGAACCTGAGAGAGCTTTAGTCGAGAGGAAACA 8700  
Db 8641 AAAAATGAAATTTATGGAATATGTAAGGAACCTGAGAGAGCTTTAGTCGAGAGGAAACA 8700  
Oy 8701 GGAAGCTGGAAGAAATTAACAAAGATCTACAGAGCAATTTGATTCAGAGAAATTTGGGA 8760  
Db 8701 GGAAGCTGGAAGAAATTAACAAAGATCTACAGAGCAATTTGATTCAGAGAAATTTGGGA 8760  
Oy 8761 ATTTGAGATTAATTTAGAAAAGATTCGAGAGATTAATGATTTGAAGATCAACGAGAGC 8820  
Db 8761 ATTTGAGATTAATTTAGAAAAGATTCGAGAGATTAATGATTTGAAGATCAACGAGAGC 8820  
Oy 8821 GGAACCAAGAGAGCTTTGTCGGAAGAAATGGAATTTCTGTGAAAAATCAATTTCAAGG 8880  
Db 8821 GGAACCAAGAGAGCTTTGTCGGAAGAAATGGAATTTCTGTGAAAAATCAATTTCAAGG 8880  
Oy 8881 GAAACCAAAATCATCTTGAATGAAGTAAGCAGAAATTTGGAACCGGAAGTGAATGTA 8940  
Db 8881 GAAACCAAAATCATCTTGAATGAAGTAAGCAGAAATTTGGAACCGGAAGTGAATGTA 8940  
Oy 8941 GATGCTTTGAATGAATTTGATGATCTTCAAGGAAAAAGTGTGCTATGCTGGAAT 9000  
Db 8941 GATGCTTTGAATGAATTTGATGATCTTCAAGGAAAAAGTGTGCTATGCTGGAAT 9000

Qy	9001	GGTATTGGAAATGTTGATAGTAAATTAATGTGTTTAAGAAAAATGTNAGAACCTAAATTCGGA	9060		
Db	9001	GGTATTGGAAATGTTGATAGTAAATTAATGTGTTTAAGAAAAATGTNAGAACCTAAATTCGGA	9060		
Qy	9061	AGACATGCTATTGTAGAACTACTGGAACAAACAAATATCAAGATTTTACAAGAGCAAAA	9120		
Db	9061	AGACATGCTATTGTAGAACTACTGGAACAAACAAATATCAAGATTTTACAAGAGCAAAA	9120		
Qy	9121	GTAAATATTCTTTGGAAAAAGAGACGCTGACCTGACCTGCATATATCGAATGTACACATT	9180		
Db	9121	GTAAATATTCTTTGGAAAAAGAGACGCTGACCTGACCTGCATATATCGAATGTACACATT	9180		
Qy	9181	TCCAATGATGATATTTAAAAATTTGGCCAAGCAGTATGCAATCTTCCATTTAATATACC	9240		
Db	9181	TCCAATGATGATATTTAAAAATTTGGCCAAGCAGTATGCAATCTTCCATTTAATATACC	9240		
Qy	9241	AAAAATTCAAAAAATATATATTACTTTAGCATCAAGTATGATGCAATGTGAATGTTTCAT	9300		
Db	9241	AAAAATTCAAAAAATATATATTACTTTAGCATCAAGTATGATGCAATGTGAATGTTTCAT	9300		
Qy	9301	GGGGTGGCTGAAGCAAGAGGTGTCAGAGACCCAAAGCAGTTAGTGTNAGAAATCAATA	9360		
Db	9301	GGGGTGGCTGAAGCAAGAGGTGTCAGAGACCCAAAGCAGTTAGTGTNAGAAATCAATA	9360		
Qy	9361	AATAGAACTAATTAATGTGATTTTAGCAGAAAAATTTAAACAGAGGGGAAACATCAATGTA	9420		
Db	9361	AATAGAACTAATTAATGTGATTTTAGCAGAAAAATTTAAACAGAGGGGAAACATCAATGTA	9420		
Qy	9421	TATGCGGATATGATTAATAAATTTATATATTAAGTAAGAACAATTTCAAGCTATTGCGGAT	9480		
Db	9421	TATGCGGATATGATTAATAAATTTATATATTAAGTAAGAACAATTTCAAGCTATTGCGGAT	9480		
Qy	9481	GCCAAAGTCATGCTGCAAGCTGCTTCGGCACTGCACTATTGAAAAAATGAAAGTAAA	9540		
Db	9481	GCCAAAGTCATGCTGCAAGCTGCTTCGGCACTGCACTATTGAAAAAATGAAAGTAAA	9540		
Qy	9541	TTTAATTAATGCGATCCGAGAAATTTTAAATATATCTGCAAGATTGGAAGGAAAGCTAAT	9600		
Db	9541	TTTAATTAATGCGATCCGAGAAATTTTAAATATATCTGCAAGATTGGAAGGAAAGCTAAT	9600		
Qy	9601	AAAAAACGTCGGTNGAGTCTTAATCAGGTAGACTGTTATACGATTAATATATACATGCGAT	9660		
Db	9601	AAAAAACGTCGGTNGAGTCTTAATCAGGTAGACTGTTATACGATTAATATATACATGCGAT	9660		
Qy	9661	TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATCAATCAAAAGAGAGGAAAAAGGAAA	9720		
Db	9661	TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATCAATCAAAAGAGAGGAAAAAGGAAA	9720		
Qy	9721	AAATGA 9726			
Db	9721	AAATGA 9726			
RESULT 2					
AR442027		11130 bp	DNA	linear	PAT 20-FEB-2004
LOCUS	AR442027				
DEFINITION	Sequence 15 from patent US 6669940.				
ACCESSION	AR442027				
VERSION	AR442027.1				
KEYWORDS					
SOURCE					
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 11130)				
TITLE	Nagaraja,T.G., Stewart,G.C., Narayanan,S.K. and Chengappa,M.M.				
JOURNAL	Recombinant fusoBacterium necrophorum leukotoxin vaccine and				
FEATURES	Patent: US 6669940-A 15 30-DEC-2003;				
FEATURES	Kanasa University Research Foundation; Manhattan, KS				
source	location/Qualifiers				
	1..11130				
	/organism="unknown"				

Query Match	100.0%	Score 9726	DB 6	Length 11130
Best Local Similarity	100.0%	Pred. No. 0		
Matches 9726	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	ATGAGCGGCATCAAAAAATTAACGTTCCAGAGCAAGAAAGAGATATCATGATCTTAAAAA	60		
DB 1034	ATGAGCGGCATCAAAAAATTAACGTTCCAGAGCAAGAAAGAGATATCATGATCTTAAAAA	10933		
QY 61	GTTTAAATGATTTTGGGATGTTGATTAAACACTAGACGGTAGGGCTAATGATACATC	120		
DB 1094	GTTTAAATGATTTTGGGATGTTGATTAAACACTAGACGGTAGGGCTAATGATACATC	11533		
QY 121	ACCGGACTGAGAAATTTTGGAACAAAAATAGAAAAAAAAGATAAATGTTATGACATTACT	180		
DB 1154	ACCGGACTGAGAAATTTTGGAACAAAAATAGAAAAAAAAGATAAATGTTATGACATTACT	12133		
QY 181	ACAAACAAGATTCAGGGGAGAAAGCCTTTTAAAGCTTTTAAATTAATTTGCTTTAAACGA	240		
DB 1214	ACAAACAAGATTCAGGGGAGAAAGCCTTTTAAAGCTTTTAAATTAATTTGCTTTAAACGA	12733		
QY 241	AATAATATAGCAAACTATATTTTGGGAAAAAGATAGTACGGGGGTAAATATCTTTT	300		
DB 1274	AATAATATAGCAAACTATATTTTGGGAAAAAGATAGTACGGGGGTAAATATCTTTT	13333		
QY 301	AACTTTGTCAATGGAAAAATTTGAAATGATGGATTTATCAACGAAATTCGAAAAATTA	360		
DB 1334	AACTTTGTCAATGGAAAAATTTGAAATGATGGATTTATCAACGAAATTCGAAAAATTA	13933		
QY 361	ATTGGAGAAATTTATATTTTCTTAAAGCTCGGAAGGATGGCATATGAAAAAATATGAGTT	420		
DB 1394	ATTGGAGAAATTTATATTTTCTTAAAGCTCGGAAGGATGGCATATGAAAAAATATGAGTT	14533		
QY 421	ATCATAGTGGTCTTTTCAATTAATTCGAAAAAACAAGATGATTTTAAAGAGCTTGG	480		
DB 1454	ATCATAGTGGTCTTTTCAATTAATTCGAAAAAACAAGATGATTTTAAAGAGCTTGG	15133		
QY 481	GAAGAGCCAAACATGATTAATTTTAAATGAAATCAATTCAGTATGAAAAAGTAA	540		
DB 1514	GAAGAGCCAAACATGATTAATTTTAAATGAAATCAATTCAGTATGAAAAAGTAA	15733		
QY 541	ATTCATTTGAATTCGAAATGGAAGCATTCGGTAAAGAAAAATCATGCTGTTGAAGGC	600		
DB 1574	ATTCATTTGAATTCGAAATGGAAGCATTCGGTAAAGAAAAATCATGCTGTTGAAGGC	16333		
QY 601	ATCGGTTATATGCGCGGATTAATAGATGTAAGAAATCTGCAATCTAAAGACAGAAAT	660		
DB 1634	ATCGGTTATATGCGCGGATTAATAGATGTAAGAAATCTGCAATCTAAAGACAGAAAT	16933		
QY 661	ACAGATTTTAAAAATTTAGTCATATTAAGTATGCAATTAATTCGTCTGACCGAGAT	720		
DB 1694	ACAGATTTTAAAAATTTAGTCATATTAAGTATGCAATTAATTCGTCTGACCGAGAT	17533		
QY 721	TTAAAGCTACCAAGCAAAAATCTGAGATATTAATCTTCAAGCTCACAATGATTTCTCT	780		
DB 1754	TTAAAGCTACCAAGCAAAAATCTGAGATATTAATCTTCAAGCTCACAATGATTTCTCT	18133		
QY 781	CAAAAAGCTATGGGAAAAAATTTCAACTGTTGGAAAAAGAAATAGAAAGAAATATGTAAGAA	840		
DB 1814	CAAAAAGCTATGGGAAAAAATTTCAACTGTTGGAAAAAGAAATAGAAAGAAATATGTAAGAA	18733		
QY 841	AATACCAAGCAAAATATGATCTGATGCTGTATTTGAAAGCAGATGCAATATATAAAAAT	900		
DB 1874	AATACCAAGCAAAATATGATCTGATGCTGTATTTGAAAGCAGATGCAATATATAAAAAT	19333		
QY 901	AGTGGAAAGCTACCAATGAGAGATTTATTAAGAAAGAGAGGAAAAAAGAACTTATAC	960		
DB 1934	AGTGGAAAGCTACCAATGAGAGATTTATTAAGAAAGAGAGGAAAAAAGAACTTATAC	19933		
QY 961	ACTCTTTAAGTTATTCAGATGTGGAAGCTTCGTAAAGTAAATAAAGAAAAAGTCTA	1020		

Db 1994 ACTCCTTAAGTTATTCAGATGTGAGAGCTTCCGTAAAGATTAATAAGAAAGTCATA 2053  
Qy 1021 GGAAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCACTTTAGTACT 1080  
Db 2054 GGAAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCACTTTAGTACT 2113  
Qy 1081 AAGCTTGAAGACACTTTTATGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAAT 1140  
Db 2114 AAGCTTGAAGACACTTTTATGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAAT 2173  
Qy 1141 GGATTTTATGTTTATTTGACAAGTACAGTGTCTGTTATTTGAAAAGTCCAAAGTC 1200  
Db 2174 GGATTTTATGTTTATTTGACAAGTACAGTGTCTGTTATTTGAAAAGTCCAAAGTC 2233  
Qy 1201 GAAGCAACAGAGAAAGCAAAATATTCATTCTTACAGTGTGAGTAAAGCACTATGGGA 1260  
Db 2234 GAAGCAACAGAGAAAGCAAAATATTCATTCTTACAGTGTGAGTAAAGCACTATGGGA 2293  
Qy 1261 GAGGCTACTTCCCACTTAATAATTTACCAATTTATTTGGAGAAAGCCAAATGGAAAACCT 1320  
Db 2294 GAGGCTACTTCCCACTTAATAATTTACCAATTTATTTGGAGAAAGCCAAATGGAAAACCT 2353  
Qy 1321 CTCAGTATCGAGCGGATATATTTCTGCAAAAAGTAAATCCATGTAACTATTTGAAGA 1380  
Db 2354 CTCAGTATCGAGCGGATATATTTCTGCAAAAAGTAAATCCATGTAACTATTTGAAGA 2413  
Qy 1381 GAAGTAAATTCGAAGGAAAGACAGATATTAATCTTCAAAATCTGAAAATATCTATGTGCT 1440  
Db 2414 GAAGTAAATTCGAAGGAAAGACAGATATTAATCTTCAAAATCTGAAAATATCTATGTGCT 2473  
Qy 1441 TCTGTTTCTGTTGGAACGATGAGAGTTCATTAAGTACTTTCACTATTTGGTGAAG 1500  
Db 2474 TCTGTTTCTGTTGGAACGATGAGAGTTCATTAAGTACTTTCACTATTTGGTGAAG 2533  
Qy 1501 GAAGGAGAAATTAATCTTCCGTCAAGATTTGCTAAAGAGCAAAAAGTAAATCAGAAAG 1560  
Db 2534 GAAGGAGAAATTAATCTTCCGTCAAGATTTGCTAAAGAGCAAAAAGTAAATCAGAAAG 2593  
Qy 1561 GATGATGTAATGTGAGAAAGTGAAGGATTAATTCATTGAGCTGCTGTAAGAGTGA 1620  
Db 2594 GATGATGTAATGTGAGAAAGTGAAGGATTAATTCATTGAGCTGCTGTAAGAGTGA 2653  
Qy 1621 TTGGGGGATAGTGTAAATGGGGTTGGGCTGCAAAATTTCTTAATCTTAATGCTCTCC 1680  
Db 2654 TTGGGGGATAGTGTAAATGGGGTTGGGCTGCAAAATTTCTTAATCTTAATGCTCTCC 2713  
Qy 1681 CGATATGATGTAGATGATATCTTACATGCCAAGAGCACTAAATGTGAGGCTCATTAAC 1740  
Db 2714 CGATATGATGTAGATGATATCTTACATGCCAAGAGCACTAAATGTGAGGCTCATTAAC 2773  
Qy 1741 ATTACTTAAATAATAGTGTCTGCAAAAGATCTGATTTGGGAATCTTCAAGTTTATGAAT 1800  
Db 2774 ATTACTTAAATAATAGTGTCTGCAAAAGATCTGATTTGGGAATCTTCAAGTTTATGAAT 2833  
Qy 1801 GATCAGCTTTATGAATCAGGTCATCTTAATCAATTTTAAATGATGCAATTAAGACGGGTT 1860  
Db 2834 GATCAGCTTTATGAATCAGGTCATCTTAATCAATTTTAAATGATGCAATTAAGACGGGTT 2893  
Qy 1861 GAGGAGACAGTGTCAATGAGAAATTAAGATTAAGTAAAGCACTAATTTAGTGTGCT 1920  
Db 2894 GAGGAGACAGTGTCAATGAGAAATTAAGATTAAGTAAAGCACTAATTTAGTGTGCT 2953  
Qy 1921 GTGTCTGCAACATATAGCAAAATCATTAATTTCTGCTTCTGTGCAATATGAGAGAGTGA 1980  
Db 2954 GTGTCTGCAACATATAGCAAAATCATTAATTTCTGCTTCTGTGCAATATGAGAGAGTGA 3013  
Qy 1981 AGACTTTTCTTCAAGAGTGTGAAGGAGATTAATGTAAGGCACTTAATGAAGCTCAAAATCTT 2040  
Db 3014 AGACTTTTCTTCAAGAGTGTGAAGGAGATTAATGTAAGGCACTTAATGAAGCTCAAAATCTT 3073  
Qy 2041 CGAGCGACTAGTCAAGTGAAGTGTGCTGTACGAAGAGAAAGAAAGAACTTAT 2100  
Db 3074 CGAGCGACTAGTCAAGTGAAGTGTGCTGTACGAAGAGAAAGAAAGAACTTAT 3133

Qy 2101 GGAAATGACAGGTTTTTTTATGAAAACTATTAATAATATGCTTGTGACAAATGCCAT 2160  
Db 3134 GGAAATGACAGGTTTTTTTATGAAAACTATTAATAATATGCTTGTGACAAATGCCAT 3193  
Qy 2161 CATGTGAATGTTGATCGGAAGAAAAATGATATCAACAGTGAATTAATTAATGATAT 2220  
Db 3194 CATGTGAATGTTGATCGGAAGAAAAATGATATCAACAGTGAATTAATTAATGATAT 3253  
Qy 2221 AAAAATCCTTCAAAAATGCGAAAGTCTGTTATGTAATAATAGAACTTTAAAGAGCT 2280  
Db 3254 AAAAATCCTTCAAAAATGCGAAAGTCTGTTATGTAATAATAGAACTTTAAAGAGCT 3313  
Qy 2281 TTGGAAAAAGAAAGAAATCCAGAAATATATCCGAAAGATATTAATCTATTTGAAAA 2340  
Db 3314 TTGGAAAAAGAAAGAAATCCAGAAATATATCCGAAAGATATTAATCTATTTGAAAA 3373  
Qy 2341 TTATTTGAATGCAATTTTCAGAAAAATTTGATGAAAAACCGAGCTTTTACTAATGTGAA 2400  
Db 3374 TTATTTGAATGCAATTTTTCAGAAAAATTTGATGAAAAACCGAGCTTTTACTAATGTGAA 3433  
Qy 2401 AGAATGCAATTAATTTCTCCGATGGAATCTTAACAAACAGGAATGCTATTAAGATGCA 2460  
Db 3434 AGAATGCAATTAATTTCTCCGATGGAATCTTAACAAACAGGAATGCTATTAAGATGCA 3493  
Qy 2461 AACTATGTTCCAGGAGAAATGAAAAATTTAGAGAAAAATTCAGAAAGATTTTAAAGCT 2520  
Db 3494 AACTATGTTCCAGGAGAAATGAAAAATTTAGAGAAAAATTTACGAAAGATTTTAAAGCT 3553  
Qy 2521 TTTTCAGAAAGATGATGATGCACTGATTAAGAAATTTGAAATTTTACAGAGTAGAAAT 2580  
Db 3554 TTTTCAGAAAGATGATGATGCACTGATTAAGAAATTTGAAATTTTACAGAGTAGAAAT 3613  
Qy 2581 TATGCAAAATTTTCAACTTTTACCTTCCGAGCTAATGAGAAAGAGATGTTCTTCT 2640  
Db 3614 TATGCAAAATTTTCAACTTTTACCTTCCGAGCTAATGAGAAAGAGATGTTCTTCT 3673  
Qy 2641 GTGGAGAGAGCTGTTTCCGTGGGTAGAACAGAGAAATTAAGCAAGGATCCGTTGAAAA 2700  
Db 3674 GTGGAGAGAGCTGTTTCCGTGGGTAGAACAGAGAAATTAAGCAAGGATCCGTTGAAAA 3733  
Qy 2701 GGAGCTTAACTTGTGCAAAAAAAGATTTTAAATTAATAAGCTATCAATTAAGCAGAAAC 2760  
Db 3734 GGAGCTTAACTTGTGCAAAAAAAGATTTTAAATTAATAAGCTATCAATTAAGCAGAAAC 3793  
Qy 2761 GTGAATTTAGTTGAAATATTTGACCTTGGAGAGACATATCCGAAAGTGCATCCGA 2820  
Db 3794 GTGAATTTAGTTGAAATATTTGACCTTGGAGAGACATATCCGAAAGTGCATCCGA 3853  
Qy 2821 GGAAGTTAAATGTTCAAGAGTCGAAAAATTCAGCTATGCTGAAGCTTAAAGAAAAAGCT 2880  
Db 3854 GGAAGTTAAATGTTCAAGAGTCGAAAAATTCAGCTATGCTGAAGCTTAAAGAAAAAGCT 3913  
Qy 2881 GAATTAATCAGAGAAAAATTAATGAGATGCAATTTGAAAGCACTTTTCAATGAGCGGA 2940  
Db 3914 GAATTAATCAGAGAAAAATTAATGAGATGCAATTTGAAAGCACTTTTCAATGAGCGGA 3973  
Qy 2941 TCTTTTAAATGTGGCTCAGGTGGGAATGCAATCAATGGAATGGGAAGTTATATGTGAGGT 3000  
Db 3974 TCTTTTAAATGTGGCTCAGGTGGGAATGCAATCAATGGAATGGGAAGTTATATGTGAGGT 4033  
Qy 3001 ATCAGTAAGCAAGAGTTTCCATGTAAGCAAGCAATTTTGAAGCTAATTAATAAAATTT 3060  
Db 4034 ATCAGTAAGCAAGAGTTTCCATGTAAGCAAGCAATTTTGAAGCTAATTAATAAAATTT 4093  
Qy 3061 GCTTTAAACAGTAAGATATCTTCTGTTTGAATGCTGCGGTTCAAGCGGAATCCGA 3120  
Db 4094 GCTTTAAACAGTAAGATATCTTCTGTTTGAATGCTGCGGTTCAAGCGGAATCCGA 4153  
Qy 3121 ACGAAAAATGCGGCTCGGGTGTCTGTTGCGGTAAATGATTTATTTCAACAA 3180  
Db 4154 ACGAAAAATGCGGCTCGGGTGTCTGTTGCGGTAAATGATTTATTTCAACAA 4213

QY 3181 GCTTCATTGAAGATTAATGACGAGGACAAAGTAAATATGATTAAGATTAAGATGATA 3240  
DB 4214 GCTTCATTGAAGATTAATGACGAGGACAAAGTAAATATGATTAAGATTAAGATGATA 4273  
QY 3241 GTPAACGTAACTGCGGAATCTTTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 3300  
DB 4274 GTPAACGTAACTGCGGAATCTTTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 4333  
QY 3301 ATTTCTGTTGCGGAGGATTAATAAGTTGAAGTAAACGAGTAAACGAGTAAACGAGTAA 3360  
DB 4334 ATTTCTGTTGCGGAGGATTAATAAGTTGAAGTAAACGAGTAAACGAGTAAACGAGTAA 4393  
QY 3361 TCAGAGAAAGAACCGAGAGGATTTTTCGCAAAATCGGAACAAAGTGAATCTGTAA 3420  
DB 4394 TCAGAGAAAGAACCGAGAGGATTTTTCGCAAAATCGGAACAAAGTGAATCTGTAA 4453  
QY 3421 AATTAATAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 3480  
DB 4454 AATTAATAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 4513  
QY 3481 GAGGTAAAGAAAGCGGAGGATCTTCCTTCGAAAGTTTCATACCTCCGATTAAGGACCG 3540  
DB 4514 GAGGTAAAGAAAGCGGAGGATCTTCCTTCGAAAGTTTCATACCTCCGATTAAGGACCG 4573  
QY 3541 TCTTTCAAGTTTGGAGCTTCTGGAAGTGTTCCTTCAATATATTAATAAGAAACATCT 3600  
DB 4574 TCTTTCAAGTTTGGAGCTTCTGGAAGTGTTCCTTCAATATATTAATAAGAAACATCT 4633  
QY 3601 GCTGTCGTGATGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 3660  
DB 4634 GCTGTCGTGATGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 4693  
QY 3661 TCTTCGATTTCACTTTTGTGAGCATGAGGCGGATCTGTCGCACTTCAGTGAATCAT 3720  
DB 4694 TCTTCGATTTCACTTTTGTGAGCATGAGGCGGATCTGTCGCACTTCAGTGAATCAT 4753  
QY 3721 ATTGGAAGTGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 3780  
DB 4754 ATTGGAAGTGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 4813  
QY 3781 ATTCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 3840  
DB 4814 ATTCAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGAT 4873  
QY 3841 AAAAGTAAATGCTTTGATGAGGAACTCAAGTACGACGAGGAGGATTTGGAAGCAT 3900  
DB 4874 AAAAGTAAATGCTTTGATGAGGAACTCAAGTACGACGAGGAGGATTTGGAAGCAT 4933  
QY 3901 AAAAGTAAATGCTTTGATGAGGAACTCAAGTACGACGAGGAGGATTTGGAAGCAT 3960  
DB 4934 AAAAGTAAATGCTTTGATGAGGAACTCAAGTACGACGAGGAGGATTTGGAAGCAT 4993  
QY 3961 GTGAACAAATGAAGTTCTGCAAAATGAGAAATTAATACGATACGAGAAATCTGAAAGC 4020  
DB 4994 GTGAACAAATGAAGTTCTGCAAAATGAGAAATTAATACGATACGAGAAATCTGAAAGC 5053  
QY 4021 CAAAAAATGAGATGATGATCACTGCTATCAAGCGGACACCCAAATGACAGAGCTTTA 4080  
DB 5054 CAAAAAATGAGATGATGATCACTGCTATCAAGCGGACACCCAAATGACAGAGCTTTA 5113  
QY 4081 AATTTACAAAGCTGGAAGTCAATGGAATGTAAGGGCTACTGTGACTGTGCTCAATTA 4140  
DB 5114 AATTTACAAAGCTGGAAGTCAATGGAATGTAAGGGCTACTGTGACTGTGCTCAATTA 5173  
QY 4141 AACCAAAAGTAAATGCTTTCTATTAAGTGTGGAAGTATATCACTTAATGAGAGGAC 4200  
DB 5174 AACCAAAAGTAAATGCTTTCTATTAAGTGTGGAAGTATATCACTTAATGAGAGGAC 5233  
QY 4201 GCAAAAGCTCTTTTACCAACCACTCAAGTGAATGAGTGAAGGAGGAGGAGCAAT 4260  
DB 5234 GCAAAAGCTCTTTTACCAACCACTCAAGTGAATGAGTGAAGGAGGAGGAGCAAT 5293  
QY 4261 AGTTCTGAGCGGATTAAGAAATTAATCAAGGGCTGTTCTGTCAATTAAGATTAAGAT 4320

DB 5294 AGTTCTGAGCGGATTAAGAAATTAATCAAGGGCTGTTCTGTCAATTAAGATTAAGAT 5353  
QY 4321 GACGTGAAGGCTAGGCTTAATTAATCTTCATTCGAGAGGCTTAATGAATGAATGTCAT 4380  
DB 5354 GACGTGAAGGCTAGGCTTAATTAATCTTCATTCGAGAGGCTTAATGAATGAATGTCAT 5413  
QY 4381 GCCAAAGATGTCAAAGAGGCTTCTGATCTACCAAAAGATTAAGAGCTTTACTAAATGA 4440  
DB 5414 GCCAAAGATGTCAAAGAGGCTTCTGATCTACCAAAAGATTAAGAGCTTTACTAAATGA 5473  
QY 4441 AAAGTAAATTAATTAATTAAGATTCGTGTATTAATACGACTGGAATGTTTATTAATACG 4500  
DB 5474 AAAGTAAATTAATTAATTAAGATTCGTGTATTAATACGACTGGAATGTTTATTAATACG 5533  
QY 4501 AAAGTAAATTAATTAATTAAGATTCGTGTATTAATACGACTGGAATGTTTATTAATACG 4560  
DB 5534 AAAGTAAATTAATTAATTAAGATTCGTGTATTAATACGACTGGAATGTTTATTAATACG 5593  
QY 4561 TCGGTTGCTGGAACGATTAATTCGCTGGAAGGATGATTAAGAGTCAATTAAGATTAAGAT 4620  
DB 5594 TCGGTTGCTGGAACGATTAATTCGCTGGAAGGATGATTAAGAGTCAATTAAGATTAAGAT 5653  
QY 4621 AATTAATTAATTAAGAGATTTGAAGTGAAGCAATTAAGAGGCGGAGAGATTAATTAAT 4680  
DB 5654 AATTAATTAATTAAGAGATTTGAAGTGAAGCAATTAAGAGGCGGAGAGATTAATTAAT 5713  
QY 4681 GCGAAACATGTAATGTAAGAGGCAAAATCATCTACTGTGTTGTAAGGCGGCTCTGGA 4740  
DB 5714 GCGAAACATGTAATGTAAGAGGCAAAATCATCTACTGTGTTGTAAGGCGGCTCTGGA 5773  
QY 4741 CTGCTATCAACAAAGATGCTTTTTCAGAAATGGAATCTGAGCATGAGCAATGCAACTATCA 4800  
DB 5774 CTGCTATCAACAAAGATGCTTTTTCAGAAATGGAATCTGAGCATGAGCAATGCAACTATCA 5833  
QY 4801 AATGACATGATTCGAAAGTGAATTAAGAGAAATTTCTGCTGATTCCTTAATGTTGAAC 4860  
DB 5834 AATGACATGATTCGAAAGTGAATTAAGAGAAATTTCTGCTGATTCCTTAATGTTGAAC 5893  
QY 4861 GCAAAATTAATTCATCTTGGGGGTGAATGTTGCGGGAACCATGCGGCTCTCTTCAGC 4920  
DB 5894 GCAAAATTAATTCATCTTGGGGGTGAATGTTGCGGGAACCATGCGGCTCTCTTCAGC 5953  
QY 4921 GCGGTAGAGGCTGCTTTTTCGAAATTAATCTCTTCAATTAATAAATCTCTGTTGATTA 4980  
DB 5954 GCGGTAGAGGCTGCTTTTTCGAAATTAATCTCTTCAATTAATAAATCTCTGTTGATTA 6013  
QY 4981 GGAACGAAGTAAATCTTTTGAAGAAATTAATCAAAAGTCAATGTAAGAGCTTTGAAT 5040  
DB 6014 GGAACGAAGTAAATCTTTTGAAGAAATTAATCAAAAGTCAATGTAAGAGCTTTGAAT 6073  
QY 5041 GATTCATTAATTAACAAAGTTCTGCTGGAAGGCTGCAAGTATTAAGAGGCTGGAATC 5100  
DB 6074 GATTCATTAATTAACAAAGTTCTGCTGGAAGGCTGCAAGTATTAAGAGGCTGGAATC 6133  
QY 5101 GAGGAAATGATATCTGCAATCTGTGTTCTGATGAACGGAAGCTTTAGTGAATCT 5160  
DB 6134 GAGGAAATGATATCTGCAATCTGTGTTCTGATGAACGGAAGCTTTAGTGAATCT 6193  
QY 5161 GAGTTGAAGAGTAAAGTTCTTTCAATGATGACAAAGATCAAAAAATTAATTAATCA 5220  
DB 6194 GAGTTGAAGAGTAAAGTTCTTTCAATGATGACAAAGATCAAAAAATTAATTAATCA 6253  
QY 5221 ATTGCGGAATTAAGCAATGGAAGAAAGCGGTGAGTTGAGCAACAGTTGCTCATCA 5280  
DB 6254 ATTGCGGAATTAAGCAATGGAAGAAAGCGGTGAGTTGAGCAACAGTTGCTCATCA 6313  
QY 5281 AATATTGAAAAATCAATCAATTAATGATTAATGATTAATTAATTAATTAATTAATTAAT 5340  
DB 6314 AATATTGAAAAATCAATCAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 6373  
QY 5341 GATCAAGTAAAGAAATTAATCAATGATGATGCAAAAGATTAATTAATTAATTAATTAAT 5400

Db 6374 GATCAAGATAGAAAAATATCAATGTGACTGCAAAAAGATTATATATGCAACATATCTATA 6433  
Qy 5401 GCAAGTCGAGTTGGAGAGCAAAAGAGCTCTGGCAAGAGCTTTCGCAAGTACTACC 5460  
Db 6434 GCAAGTCGAGTTGGAGAGCAAAAGAGCTCTGGCAAGAGCTTTCGCAAGTACTACC 6493  
Qy 5461 TTGAATTAAGACAGTTTCTTCTCATGTGTGATCAAACTGATATTAAGCAAAAGATTTAAGNA 5520  
Db 6494 TTGAATTAAGACAGTTTCTTCTCATGTGTGATCAAACTGATATTAAGCAAAAGATTTAAGNA 6553  
Qy 5521 GAAAAATATGAAATTAAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACAGTCA 5580  
Db 6554 GAAAAATATGAAATTAAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACAGTCA 6613  
Qy 5581 GTGGTCACAAATCCGACAGTCTTTCGAGCAAGTGGACAAGCTGCAATAGAGACTGGA 5640  
Db 6614 GTGGTCACAAATCCGACAGTCTTTCGAGCAAGTGGACAAGCTGCAATAGAGACTGGA 6673  
Qy 5641 GTAGCAGTTAATTAATTAACAAAAATCTTGCACATATAAAAATAGTACTCAAAAT 5700  
Db 6674 GTAGCAGTTAATTAATTAACAAAAATCTTGCACATATAAAAATAGTACTCAAAAT 6733  
Qy 5701 GTACGAAATGCTTGTGTAAAAAGCAAAATCTCATCTTATTAACAAATGGAATTTGA 5760  
Db 6734 GTACGAAATGCTTGTGTAAAAAGCAAAATCTCATCTTATTAACAAATGGAATTTGA 6793  
Qy 5761 GCTGAGTTGGAGCTGGAGAGCTGGAGTGCAGGTTCTGTAGCACTGAAATTAAGATTGA 5820  
Db 6794 GCTGAGTTGGAGCTGGAGAGCTGGAGTGCAGGTTCTGTAGCACTGAAATTAAGATTGA 6853  
Qy 5821 AATTAATGATGCGAATTAATCATGCAAAAAATCACTGCGAAGGAAATGCGAGTT 5880  
Db 6854 AATTAATGATGCGAATTAATCATGCAAAAAATCACTGCGAAGGAAATGCGAGTT 6913  
Qy 5881 ATTACAGAGTCTGATCGGTAATGCTAAATTATGCAAGAAACAGTGTCTGAGTGGCCGT 5940  
Db 6914 ATTACAGAGTCTGATCGGTAATGCTAAATTATGCAAGAAACAGTGTCTGAGTGGCCGT 6973  
Qy 5941 GCAGCAATAGAGCTCAACCAAGTGTGAATGAAATTAACAGATCTACAAAAGATATGA 6000  
Db 6974 GCAGCAATAGAGCTCAACCAAGTGTGAATGAAATTAACAGATCTACAAAAGATATGA 7033  
Qy 6001 AAAAGATTCAAGTGTGCTAAGAAAGAAAGATTAATTAATTAATTAATTAATTAATTAAT 6060  
Db 7034 AAAAGATTCAAGTGTGCTAAGAAAGAAAGATTAATTAATTAATTAATTAATTAATTAAT 7093  
Qy 6061 GTAGATTAAGTGTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6120  
Db 7094 GTAGATTAAGTGTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7153  
Qy 6121 AAAAGAAAAATTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6180  
Db 7154 AAAAGAAAAATTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7213  
Qy 6181 AAATCTTTTATGGCAAAATGCGCTGTTCAAGCAAGCCGAGATGGCAGAAATCTTAAT 6240  
Db 7214 AAATCTTTTATGGCAAAATGCGCTGTTCAAGCAAGCCGAGATGGCAGAAATCTTAAT 7273  
Qy 6241 ATCAACAAGTTTATGAGAAACAGAAAGCTCTGTGAAAAATCTATATTAATTAATTAATTAAT 6300  
Db 7274 ATCAACAAGTTTATGAGAAACAGAAAGCTCTGTGAAAAATCTATATTAATTAATTAATTAAT 7333  
Qy 6301 CATTATTTCTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6360  
Db 7334 CATTATTTCTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7393  
Qy 6361 GTTGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6420  
Db 7394 GTTGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7453  
Qy 6421 AAGACAAAGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6480  
Db 7454 AAGACAAAGTGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7513

Qy 6481 ACACAGATTTCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTGCAGACGCGGGTGA 6540  
Db 7514 ACACAGATTTCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTGCAGACGCGGGTGA 7573  
Qy 6541 GAGAGCGAGTGGCAGAAACCGTTTCCGTAAATCAATTTGCAAGAAACCGAATAGAT 6600  
Db 7574 GAGAGCGAGTGGCAGAAACCGTTTCCGTAAATCAATTTGCAAGAAACCGAATAGAT 7633  
Qy 6601 GTGGAAGAAAGAAATTTTGGTAAATAAAGCTGATTAACAGCAAAAGTTATAGTTCT 6660  
Db 7634 GTGGAAGAAAGAAATTTTGGTAAATAAAGCTGATTAACAGCAAAAGTTATAGTTCT 7693  
Qy 6661 GTTGCAATTTGAAATGCGCAGTCCGAGTGGCTGCAAAAGAGCTGGAATTTGACACAGA 6720  
Db 7694 GTTGCAATTTGAAATGCGCAGTCCGAGTGGCTGCAAAAGAGCTGGAATTTGACACAGA 7753  
Qy 6721 GTGGCAGTTACCAAAAGATGAATCAAAACAGAGCAAGATGAATAATTTCTAAATTAATG 6780  
Db 7754 GTGGCAGTTACCAAAAGATGAATCAAAACAGAGCAAGATGAATAATTTCTAAATTAATG 7813  
Qy 6781 ACTGCAAAACAGTTGATGATTAATTAACAGAAAAATGATTAATTAATTAATTAATTAATTAAT 6840  
Db 7814 ACTGCAAAACAGTTGATGATTAATTAACAGAAAAATGATTAATTAATTAATTAATTAATTAAT 7873  
Qy 6841 TCAGCCGAGCTGGAATTTCTGACGCGAGATCTGAGTGGTCTGCTCAATATATAT 6900  
Db 7874 TCAGCCGAGCTGGAATTTCTGACGCGAGATCTGAGTGGTCTGCTCAATATATAT 7933  
Qy 6901 GCAAAATTAAGTGAAGAAACAGATATGATATGATATGATATGATATGATATGATATGATATGAT 6960  
Db 7934 GCAAAATTAAGTGAAGAAACAGATATGATATGATATGATATGATATGATATGATATGATATGAT 7993  
Qy 6961 GTAAAAAGCTTTAATTAATTTTCCAAATTTCTTGAACAGCCGGTGGAGAGCCGAGTCTT 7020  
Db 7994 GTAAAAAGCTTTAATTAATTTTCCAAATTTCTTGAACAGCCGGTGGAGAGCCGAGTCTT 8053  
Qy 7021 GCAGCAGTTACCGAGTGGTCTTCTGTAAACATTAATTAATTAATTAATTAATTAATTAATTAAT 7080  
Db 8054 GCAGCAGTTACCGAGTGGTCTTCTGTAAACATTAATTAATTAATTAATTAATTAATTAATTAAT 8113  
Qy 7081 CAAATTAATCTGATTTGAATTTCCGTAACAGAAAAATGAATGAATGCAAAAGAGAA 7140  
Db 8114 CAAATTAATCTGATTTGAATTTCCGTAACAGAAAAATGAATGAATGCAAAAGAGAA 8173  
Qy 7141 AAAAATTAATTAACAAACAGCAGCAATGCAAGAAATGCAAGAAATGCAAGAAATGCAAGAAAT 7200  
Db 8174 AAAAATTAATTAACAAACAGCAGCAATGCAAGAAATGCAAGAAATGCAAGAAATGCAAGAAAT 8233  
Qy 7201 GTCTTGTAAATTAATTTTGGAAACAGCTGTAAGAAATTAATTAATTAATTAATTAATTAAT 7260  
Db 8234 GTCTTGTAAATTAATTTTGGAAACAGCTGTAAGAAATTAATTAATTAATTAATTAATTAATTAAT 8293  
Qy 7261 ACAGAAAGTTTAAATTAATTTAGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7320  
Db 8294 ACAGAAAGTTTAAATTAATTTAGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8353  
Qy 7321 GCTACGAAAAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7380  
Db 8354 GCTACGAAAAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8413  
Qy 7381 GATAGAGGATATCTACAGGAGAAAGAAATTAAGCAATTTGTAAGAACTCTGATATTAAT 7440  
Db 8414 GATAGAGGATATCTACAGGAGAAAGAAATTAAGCAATTTGTAAGAACTCTGATATTAATTAAT 8473  
Qy 7441 GGAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7500  
Db 8474 GGAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8533  
Qy 7501 GGAATGCAAGTCTTCTTCCGATACAGAAACAGTGGCAGTTTCAATTAATTAATTAATTAATTAAT 7560  
Db 8534 GGAATGCAAGTCTTCTTCCGATACAGAAACAGTGGCAGTTTCAATTAATTAATTAATTAATTAAT 8593

QY 7561 TCCGAGATTACTGTTGAAAAATTCCTTTGTGAAAGCAGCTGAAAAAGTAATGTTAGATCG 7620  
Db 8594 TCCGAGATTACTGTTGAAAAATTCCTTTGTGAAAGCAGCTGAAAAAGTAATGTTAGATCG 8653  
QY 7621 GATATTACAGGAAATGTTGCTTTAAACAGATATCAAGCTCTGTAGAGCATTTGGGAATA 7680  
Db 8654 GATATTACAGGAAATGTTGCTTTAAACAGATATCAAGCTCTGTAGAGCATTTGGGAATA 8713  
QY 7681 GGAGCTGCTTACAGAAATTTAAATTTCTAATGGAAGATCAAAATATCAGTATTTAAAAATCT 7740  
Db 8714 GGAAGCTGCTTACAGAAATTTAAATTTCTAATGGAAGATCAAAATATCAGTATTTAAAAATCT 8773  
QY 7741 AAGCTATTGAAAAAATAATGATGTTATTGTTAAAGATTAATTCGGAATTGAGCGGAA 7800  
Db 8774 AAGCTATTGAAAAAATAATGATGTTATTGTTAAAGATTAATTCGGAATTGAGCGGAA 8833  
QY 7801 GCAAAAAGATTAAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAAAGAT 7860  
Db 8834 GCAAAAAGATTAAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAAAGAT 8893  
QY 7861 GAAATGAATTCAGAGGTTGAAAATTGAGAAGATTTTCAATGAGAAAATGAGTAACT 7920  
Db 8894 GAAATGAATTCAGAGGTTGAAAATTGAGAAGATTTTCAATGAGAAAATGAGTAACT 8953  
QY 7921 AGCCCTTCTAAAGAAATTGAAAGAAATCAAATGTCAAAAGTGGAAAAAGAAAAACAGAGTG 7980  
Db 8954 AGCCCTTCTAAAGAAATTGAAAGAAATCAAATGTCAAAAGTGGAAAAAGAAAAACAGAGTG 9013  
QY 7981 ACTGCTGAATCTCAAGAGACTTCTGTAGAGACAGTACGAGGGGACAGAAATTTTCCGAA 8040  
Db 9014 ACTGCTGAATCTCAAGAGACTTCTGTAGAGACAGTACGAGGGGACAGAAATTTTCCGAA 9073  
QY 8041 GCAAAAAGATGCGGAAAGCTCTTATTGAAAGTTAGTACAAAATCCGGAAGAGTATTTT 8100  
Db 9074 GCAAAAAGATGCGGAAAGCTCTTATTGAAAGTTAGTACAAAATCCGGAAGAGTATTTT 9133  
QY 8101 CATGCAGATTAATGTGAATAATGGAAGCAACATAAAATGAAAAGTAAACAGAGTTCTAA 8160  
Db 9134 CATGCAGATTAATGTGAATAATGGAAGCAACATAAAATGAAAAGTAAACAGAGTTCTAA 9193  
QY 8161 GCAGTAAACAGGTTCTGTATTGGAGAGAGTTGAGTACCAAGGACAGAGCTACTGCTGCA 8220  
Db 9194 GCAGTAAACAGGTTCTGTATTGGAGAGAGTTGAGTACCAAGGACAGAGCTACTGCTGCA 9253  
QY 8221 GGTAAACATATGATAGTAGAGTTGAGAGAAATTTGTCAGAACAAATGATTTAGTGA 8280  
Db 9254 GGTAAACATATGATAGTAGAGTTGAGAGAAATTTGTCAGAACAAATGATTTAGTGA 9313  
QY 8281 ATTTCTAAAGTAGAAGGTTGATGTAAGATAAAGTAACTGCTAAATCTTCTGATATCA 8340  
Db 9314 ATTTCTAAAGTAGAAGGTTGATGTAAGATAAAGTAACTGCTAAATCTTCTGATATCA 9373  
QY 8341 GGAATATGAGAGAGAAATGCGGAGCAGAGTGAATATTCTTACAGCACAAAGTAACT 8400  
Db 9374 GGAATATGAGAGAGAAATGCGGAGCAGAGTGAATATTCTTACAGCACAAAGTAACT 9433  
QY 8401 GAATCGTAGTTCGTTTACGAAAGCAAGTTAATGAATATGATTTACCAAAAAATAT 8460  
Db 9434 GAATCGTAGTTCGTTTACGAAAGCAAGTTAATGAATATGATTTACCAAAAAATAT 9493  
QY 8461 ATTTCAAGAGTCAATGCTCTTGTCTTTAAATGATACAAAGATGAGCGAATAGATCT 8520  
Db 9494 ATTTCAAGAGTCAATGCTCTTGTCTTTAAATGATACAAAGATGAGCGAATAGATCT 9553  
QY 8521 TTAGCGGTAGCGGTGTGATGCAACAAGAACAAACAAATTTTACGAGATCAACAAG 8580  
Db 9554 TTAGCGGTAGCGGTGTGATGCAACAAGAACAAACAAATTTTACGAGATCAACAAG 9613  
QY 8581 TTAACCTTACAACTGTAATGAGAGAAAGTATCTCAACTTCGTCGCAAAAGCTTTGGCT 8640  
Db 9614 TTAACCTTACAACTGTAATGAGAGAAAGTATCTCAACTTCGTCGCAAAAGCTTTGGCT 9673  
QY 8641 AAAAAATGAAAAATTAATGAAAAATGTAAGAACTGAGAGAGCTTATGTCGAGCGGAAACA 8700

Db 9674 AAAAAATGAAAAATTAATGAAAAATGTAAGAACTGAGAGAGCTTATGTCGAGCGGAAACA 9733  
QY 8701 GCAGCCGTTGAAATTTATACAAAGATCTACAGAGCATTTGGTTGACAGAAATTTGGAA 8760  
Db 9734 GCAGCCGTTGAAATTTATACAAAGATCTACAGAGCATTTGGTTGACAGAAATTTGGAA 9793  
QY 8761 ATTGAGATTAATTTAGAAACGATTGCAAGATATAATGATTTGAAAGTCAACGAGAC 8820  
Db 9794 ATTGAGATTAATTTAGAAACGATTGCAAGATATAATGATTTGAAAGTCAACGAGAC 9853  
QY 8821 GGAACCAAGAGAGTCTTGTGGAAGAAATGATTTCTGTGAAAAATACATTTCAAGG 8880  
Db 9854 GGAACCAAGAGAGTCTTGTGGAAGAAATGATTTCTGTGAAAAATACATTTCAAGG 9913  
QY 8881 GAAACAAAATCATCCATGGAAGATTAACCCGAAATTTGGAACCGGAAGTAAATGTA 8940  
Db 9914 GAAACAAAATCATCCATGGAAGATTAACCCGAAATTTGGAACCGGAAGTAAATGTA 9973  
QY 8941 GATGCTTGAATGAATCTGATGATGATCTACAGAGAAAGTGTGCTATGTGGAAT 9000  
Db 9974 GATGCTTGAATGAATCTGATGATGATCTACAGAGAAAGTGTGCTATGTGGAAT 10053  
QY 9001 GGTATTGGAATTTGATGATGATTAATGATGATTTAAGAAAAATGTAAGCCAAATTCGA 9060  
Db 10034 GGTATTGGAATTTGATGATGATTAATGATGATTTAAGAAAAATGTAAGCCAAATTCGA 10093  
QY 9061 AGACATGCTATTGTAAGAACTACAGAGAAACAGAAATTCAGAACTTTACAAAGCCAAA 9120  
Db 10094 AGACATGCTATTGTAAGAACTACAGAGAAACAGAAATTCAGAACTTTACAAAGCCAAA 10153  
QY 9121 GTAATATTCTTGGAAAAAGAGACGCTCAGCTGAGCTGCAATTCGAAATGTACACATT 9180  
Db 10154 GTAATATTCTTGGAAAAAGAGACGCTCAGCTGAGCTGCAATTCGAAATGTACACATT 10213  
QY 9181 TCCAAATGATGATTAATTAATTTGGCAAGCAGTATGCAATCTTCTCAATTAATAAC 9240  
Db 10214 TCCAAATGATGATTAATTAATTTGGCAAGCAGTATGCAATCTTCTCAATTAATAAC 10273  
QY 9241 AAAAAATTCAAAAAATTAATTAATTTAGCATCAAGTATGTAATGTAATGTAATGTTGAT 9300  
Db 10274 AAAAAATTCAAAAAATTAATTAATTTAGCATCAAGTATGTAATGTAATGTAATGTTGAT 10333  
QY 9301 GGGGTGCTGAAGCAAGAGTGCAGAGACCAAGCGACAGTTAGTGTAAAGATCAATA 9360  
Db 10334 GGGGTGCTGAAGCAAGAGTGCAGAGACCAAGCGACAGTTAGTGTAAAGATCAATA 10393  
QY 9361 AATAGAACTTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAACATCAATGTA 9420  
Db 10394 AATAGAACTTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAACATCAATGTA 10453  
QY 9421 TATGCCGATATGATTAATAATTAATTAATGTAAGACAAATTTCTAAGCTTATGCCGAT 9480  
Db 10454 TATGCCGATATGATTAATAATTAATTAATGTAAGACAAATTTCTAAGCTTATGCCGAT 10513  
QY 9481 GCCAAAGTCAATGCTGACAGCTGCTCGGCAACTGCACTATTGAAAAAATGAAATGTA 9540  
Db 10514 GCCAAAGTCAATGCTGACAGCTGCTCGGCAACTGCACTATTGAAAAAATGAAATGTA 10573  
QY 9541 TTTAATTAATGCAATCCGAAATTTAAAAATACTGGCAAGATTGGAAGGAAAGCTAAT 9600  
Db 10574 TTTAATTAATGCAATCCGAAATTTAAAAATACTGGCAAGATTGGAAGGAAAGCTAAT 10633  
QY 9601 AAAAAAAGTGGTGTGATCTTAATAGGTAGATGCTATACGGAATTAATTAATCAATGCAAT 9660  
Db 10634 AAAAAAAGTGGTGTGATCTTAATAGGTAGATGCTATACGGAATTAATTAATCAATGCAAT 10693  
QY 9661 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGAAA 9720  
Db 10694 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGAAA 10753  
QY 9721 AATAGA 9726

Db	10754	AAATGA	10755
RESULT 3	AF312861	12700 bp	DNA linear BCT 22-MAR-2005
LOCUS	AF312861		
DEFINITION	Fusobacterium necrophorum Oppf (oppf) gene, partial cds; and LktB (lktb), Lkta (lktA), and LktC (lktc) genes, complete cds.		
ACCESSION	AF312861		
VERSION	AF312861.3	GI:61676922	
KEYWORDS			
SOURCE	Fusobacterium necrophorum		
ORGANISM	Fusobacterium necrophorum		
REFERENCE	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae; Fusobacterium.		
AUTHORS	1 (bases 1 to 12700)		
TITLE	Narayanan,S.K., Nagaraaja,T.G., Chengappa,M.M. and Stewart,G.C. Cloning, sequencing, and expression of the leukotoxin gene from Fusobacterium necrophorum		
JOURNAL	Infect. Immun. 69 (9), 5447-5455 (2001)		
PUBMED	11500416		
REFERENCE	2 (bases 1 to 12700)		
AUTHORS	Oelke,A.M., Nagaraaja,T.G., Wilkerson,M.J. and Stewart,G.C.		
TITLE	The leukotoxin operon of Fusobacterium necrophorum		
JOURNAL	Unpublished		
REFERENCE	3 (bases 1 to 12700)		
AUTHORS	Narayanan,S., Stewart,G.C., Nagaraaja,T.G. and Chengappa,M.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-OCT-2000) Diagnostic Medicine / Pathobiology, Kansas State University, 1800 Denison Ave, Manhattan, KS 66506, USA		
REFERENCE	4 (bases 1 to 12700)		
AUTHORS	Stewart,G.C., Narayanan,S., Oelke,A.M., Nagaraaja,T.G. and Chengappa,M.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-FEB-2004) Diagnostic Medicine / Pathobiology, Kansas State University, 1800 Denison Ave, Manhattan, KS 66506, USA		
REFERENCE	5 (bases 1 to 12700)		
AUTHORS	Stewart,G.C., Narayanan,S., Oelke,A.M., Nagaraaja,T.G. and Chengappa,M.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-MAR-2005) Diagnostic Medicine / Pathobiology, Kansas State University, 1800 Denison Ave, Manhattan, KS 66506, USA		
REFERENCE	Sequence update by submitter		
AUTHORS	On Mar 22, 2005 this sequence version replaced gi:42759853.		
REMARK	location/Qualifiers		
COMMENT	1..12700		
FEATURES	/organism="Fusobacterium necrophorum"		
source	/mol_type="genomic DNA"		
	/strain="A25"		
	/db_xref="taxon:859"		
	<1..211		
gene	/gene="oppf"		
	<1..211		
CDS	/gene="oppf"		
	/note="ABC transporter; nucleotide binding/ATPase protein"		
	/codon_start=2		
	/transl_table=11		
	/product="Oppf"		
	/protein_id="AAS45303.1"		
gene	/db_xref="GI:42759855"		
	/translation="DLPEKIKKERLSCFLFISHDLFVVRNICDRVLMLKGIIEGK		
	KEEIPKAPKKEYTQFLLEVSMSPF"		
	685..2337		
	/gene="lktB"		
	/note="ups"		
CDS	685..2337		
	/gene="lktB"		
	/note="putative transport protein"		
	/codon_start=1		
	/transl_table=11		
	/product="LktB"		
	/protein_id="AAK27340.2"		

```

/db_xref="GI:47759854"
/translation="MFHKLKQTEFVFWFLLILSCPLLIAEGTISQSTESDAGFLMLDRN
NRIPOOGLQYLOKLTIEKQVYKTIYQKINELVYKQKTYLVCAYLPAQKLONGVLLILIE
EIVRIAMEYVLEKEVTVADDTYKTIYQKINELVYKQKTYLVCAYLPAQKLONGVLLILIE
EKTGQIITIIQGNSTSEKTYIREIIPLEKDSVSPFKELDRSLTRENLINDSEIQINMASG
KYPGTADSVFVYIEBERQOQFVPAVNDLQKNTQKGLMELWJLINSVYGNRNDLSLVS
VTEGTPSTFVPTFVPSKGGTISLQHSYGLKHLIGALAKHKTITGNSYSGVGIYHPI
LVEHKNKELSLDMVKQKRPVLDLKLWNNLNSKTKTAGGISHYEDSVFYKQNT
KQKPIPIGSDANLWYKQKRPVLDLKLWNNLNSKTKTAGGISHYEDSVFYKQNT
YAGGAIVNRYIPENFMCAHEGVFPNALSSTLYENKGEFVYLDGASLHGEBAWQNRIPSSGF
GYKIRFLEKQNTIAVSNAPFPWKCKINSISVDSNKITYITINHEP"
2353..12078
/gene="1jcta"
2353..12078
/gene="1jcta"
/ncbi_taxonomy="1"
/codon_start=1
/translation="MSGIKNNVQRTKRIISDKKYLMLIGLLINTYVRANDITLATE
NGETIKEDKDYVDTITNTKIQENMAINSNRPLATNTNNILAIPEGKNSUGNNLFPNF
VNGKILEVDIINGIDENKICGMLYVSSSGMAVGRKYNINAGFHSIIPQODFPEKSV
EAKKQKYNVNGIIPVDGKATPIPLNPGSITVEKCIYALVBEIGIYAADIRLQDILITK
GTDPRKNVINSIDRINSIGTGLKATKTSQGIISAHIDPSOKANGNSVIGRIEISV
YKQNTKANIEDDAVLEADGNTKISAKATNGRIIPKKEGEKTYNTPLSISDVASVAV
NGKATIGKRVDTIAEKNYDYDVLVYPLKALHSPKSTYGTISIPNLQFDLSTKSSV
VYGDQAKVATGATGKANIHSYSGVATMGATASPLKTITLVEKNGLISIGAGYISA
KNSNVTIEGEVYSKQKRIATSEKNTIDASVSGVGMDSNKALSYLVEGEKSSV
KIAKAKVCESEFNDVNVREBALINSIAAVKGGISGNGVANNISYNAVSRIDVQGV
YIHAKKRLVNEAHTNRNSVLOTGSDLGTSKPMNDHYSGHKLSDILAIKORPQNS
VNEEIKNKTJLNPSPVKSATIAHNNNSAVALIGSGRLSGVCSNTRALANRONLRA
TSSGSGVAVRKEBKEKLIIONAVFYGNYNKNAVTTADHAEVSEKTIIDINSEKITEY
KQPSKAKVAKETILKELKAFGEKTIPEYDPOIISIEKLINAFSEKLDGKPELNTL
GEMRTIIPDGNSKTGTELAIVYOGEMKKEBEKIPKGAISEGSGIKETLNTYV
GYCNVNAVHTNIGSANGSRDVSOGVAVSWQNSKNSYVKGAKVLAKKDQNTKRA
INAEFTVNLVNTGLARLSTSGSAVGRALVQSKNSIYAEKKEKLESGENINADML
NLPHVAGPNSGSGGNALNGMSYSGGISKARVSDDEYVLAANKIALNSKRDTSV
NNAAGAGAGTGOAAVAVAVANDYIENKASIEDDEGQSKYDKNDDEVITYABESL
EYADAKTGTINISVAGVINKYCSKSEKPEKSEBERPEFGKIKGNVDSKNTITDPS
MOLTEKINIEYSEBGKQKGLPSNVSHFDPKQFELGAGSVSPNNIKETSAVVD
GYNALIKANKKYETVESDSTPFGAGGSAALOMNHIGSGNSINSAGLAGAAVANNIG
CYKIKNIAKNTVNEAKSSVTVYVNAISGLAISDAPEGMSGAMODLSNDPIIAYDKQ
BAGEDKIHAHTNVNEAKSSVTVYVNAISGLAISDAPEGMSGAMODLSNDPIIAYDKQ
RISADBLNVANNISILGVNAVGTIAGSLSTAVGAPAPANTLHNTKSLITGTYNPRE
KGNTKVNAVANDLSHTINVSAGAGASIKQAGIGMVSVNGSDTEKALVSDSEBGVYS
SEFVNDADOKTITNIAGNNGVGAAGVAVATNTIGKSVIAYVNSKTIYANDQDQK
KININAVAKOTYMTNITAVGAGAKASVQASASTLNTGNSVSHVDDTDIDKEEBEKE
NGKKEKANNVLAENTSOYVNTATVLTSGAAGVAQGVAVNKTONTSHAIKNTOSTO
VNRALVYKSHSHSIKITGIGAGVAGAGATGVSVAVKNYNTNITALEHAKTIKAGNV
GYTSDDAVYANTAGVSGVGAARVAGASTVNEITGSTRKAYVDQSYIAKEETDPIY
TOQOVDVNDKPKKNTINEDLSQKRIENKKEKKEKNSATHTLKSILANVAGASQQA
VAGTAVINNKYGTEALVENSILNAKSYVSGGYTNSIATHTLKSILANVAGASQQA
TNIIKENTKRTVEKTMISDGEFEEBETIADSKQSGISFGVGAAGVAGVAGVS
NORAGTEVDVBEAKLIVYKKAETIYKARSYVAGIAGNAVGAQGVAGAVYKDEBA
NTBARVNSKIMTRKLDVIAENBISAGTIGSAGAGIILAAGVSGVSVNNIANKVE
DIDHSTLHSTVDNVAKELNINSITLNGGAGAGIATGVAVTINSVARSCHGNS
DLSVSEKTVNVAKELNIKOTLANIIGGAIAGNAVNFVAGVNSKVBEGTIS
VLTDLBEVNDOKVADNAPKTLIQSGISITEDTSVADSDQTOGEBKIAVKPDSI
GKRVDTTDEKKNITSGVGLGTGSLASAGVAVTNIKNSGVTVENSFYAERKAVN
RSBITENVALTVACQPGALGICAGVAVNLSNBSNPSIKNSKTLGKNDIIVYKKA
LRKEAGLITGAGVAGVAVAGAIISEKNDAGSSTLKYSTYSGRIPHADVNMEYTHK
EKENRTVABEGVAGVAVAGAIISEKNDAGSSTLKYSTYSGRIPHADVNMEYTHK
MKTAVASVATVGLGVGVYTKBAETAAKGTNVEBEGNTETRNLNAISVEGIDEBE
KYTAKSVSVNGGGIAGADVNTAQSNTSVRLRKQDVENNDYKRTISEVALLDEK
LNDTKEANIESTLAVGVHAGVNTKATPFRNKSTIGKLEETIADNTVRLVNGDQK
NNGTQKGLYAEATLAVENYTKSTGLVAGNKTIGKLEETIADNTVRLVNGDQK
GLVKGKGIYKANTIGSETTSSIEDKAIYGTGSGVNDALNELVDVLDQKSGGYGIGI

```

gene  
CDS  
12075..12512  
/gene="lkc"  
12075..12512  
/gene="lkc"  
/codon\_start=1  
/transl\_table=11  
/product="lkc"  
/protein\_id="AA545304.1"  
/db\_xref="GI:42759856"  
/translation="MKLRSEKSEFLKNSNITCFPEREYKDELTVYRSFMEVEGON  
LPMVIMDNSITNIRVOIAPVYIKDNTSEAVLSYINELNREYKFKFYVTEADAVCL  
DSCVTSIAEEFPEMYTITLVNLIHSITHTSIPMKKIMSEK"  
ORIGIN  
Query Match 100.0%; Score 9726; DB 1; Length 12700;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGGGGATCAAAAATTAACGTTGAGAGACAAAGAAAGATTCAGATTCTAAAA 60  
DB 2353 ATGAGGGGATCAAAAATTAACGTTGAGAGACAAAGAAAGATTCAGATTCTAAAA 2412  
QY 61 GTTTTATGATTTTGGGATTTGTTGATTAACACTATGACGGTGAAGGCTAATGATCAATC 120  
DB 2413 GTTTTATGATTTTGGGATTTGTTGATTAACACTATGACGGTGAAGGCTAATGATCAATC 2472  
QY 121 ACCGCGACTGAGAAATTTGGAACAAAATAGAAAAAAGATATGTTTATGACATTACT 180  
DB 2473 ACCGCGACTGAGAAATTTGGAACAAAATAGAAAAAAGATATGTTTATGACATTACT 2532  
QY 181 ACCAAACAATTCAGAGGAGAAACGCTTTTAAACGTTTAAATGATTTGCTTTAACAGAA 240  
DB 2533 ACCAAACAATTCAGAGGAGAAACGCTTTTAAACGTTTAAATGATTTGCTTTAACAGAA 2592  
QY 241 AATTAATATAGCAAAATCTATATTTTGGGAAAAAGATAGTACGGGGTAAATATCTTTT 300  
DB 2593 AATTAATATAGCAAAATCTATATTTTGGGAAAAAGATAGTACGGGGTAAATATCTTTT 2652  
QY 301 AACTTTGTCAATGAAAAATTTGAAGTAGATGGATTTATCAACGGAATTCAGAAAAATAAA 360  
DB 2653 AACTTTGTCAATGAAAAATTTGAAGTAGATGGATTTATCAACGGAATTCAGAAAAATAAA 2712  
QY 361 ATTGAGGAAATTTATATTTCTTAAGCTGGAAAGGATGGCAGTAGAGAAAAATGAGATT 420  
DB 2713 ATTGAGGAAATTTATATTTCTTAAGCTGGAAAGGATGGCAGTAGAGAAAAATGAGATT 2772  
QY 421 ATCAATGCTGTTCTTTTCATTCATTTATTTCCAAAACAAGATGATTTTAAAGAGCTTTG 480  
DB 2773 ATCAATGCTGTTCTTTTCATTCATTTATTTCCAAAACAAGATGATTTTAAAGAGCTTTG 2832  
QY 481 GAAAGAGCCAAACATGATAAGTTTAAATGGAATCATTCAGTAGAGAAAAAGTAAAA 540  
DB 2833 GAAAGAGCCAAACATGATAAGTTTAAATGGAATCATTCAGTAGAGAAAAAGTAAAA 2892  
QY 541 ATTCCATTGAATCCGAATGGAAGCATTAACGTTAGAAAGAAAAATCAATCTGTGAAGGC 600  
DB 2893 ATTCCATTGAATCCGAATGGAAGCATTAACGTTAGAAAGAAAAATCAATCTGTGAAGGC 2952  
QY 601 ATCGGTTTATGCGCGGATATTAGATTGAAAGATGCAATACTAATAAGCAGGAATT 660  
DB 2953 ATCGGTTTATGCGCGGATATTAGATTGAAAGATGCAATACTAATAAGCAGGAATT 3012  
QY 661 ACAGATTTTAAAAATTAGTCAATATTAAGTATGATCGATAAATTCGTGTCGACCGAGAT 720  
DB 3013 ACAGATTTTAAAAATTAGTCAATATTAAGTATGATCGATAAATTCGTGTCGACCGAGAT 3072  
QY 721 TTTAAAGCTACCAAGCAAAATCTGGAGATATTATCTTTCAAGCTCACTAAGTTCTCCT 780

DB 3073 TTTAAAGCTACCAAGCAAAATCTGGAGATATTATTTCTTCAAGCTCACATGATTTCTCCT 3132  
QY 781 CAAAAGCTATGGAAAAAAATTTCAACTGTTGGAAAAAGAAATAGAAATATGTAANAAGA 840  
DB 3133 CAAAAGCTATGGAAAAAAATTTCAACTGTTGGAAAAAGAAATAGAAATATGTAANAAGA 3192  
QY 841 AATAACCAAGCAAAATATTGAATCTGATGCTGTATTGGAAAGCATGGAATATATAAAATT 900  
DB 3193 AATAACCAAGCAAAATATTGAATCTGATGCTGTATTGGAAAGCATGGAATATATAAAATT 3252  
QY 901 AGTCGAAAGCTACAAAATGGAGATTTTAAAGAAAGAGGAAAAAGAACTTATTAAC 960  
DB 3253 AGTCGAAAGCTACAAAATGGAGATTTTAAAGAAAGAGGAAAAAGAACTTATTAAC 3312  
QY 961 ACTCCTTAAGTTTATCAGATGTAAGCTCCGTAAGAGTAAATTAAGAAAGAGTCATA 1020  
DB 3313 ACTCCTTAAGTTTATCAGATGTAAGCTCCGTAAGAGTAAATTAAGAAAGAGTCATA 3372  
QY 1021 GAAAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTTAGTTACT 1080  
DB 3373 GAAAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTTAGTTACT 3432  
QY 1081 AAGCTTGAAAGCACTCTTTAGCTTTGTTACAGGTTCTATTTCTCTATGCAATTTAAAT 1140  
DB 3433 AAGCTTGAAAGCACTCTTTAGCTTTGTTACAGGTTCTATTTCTCTATGCAATTTAAAT 3492  
QY 1141 GGAATTTAGGTTTATGACAAGTAACTCAAGTGTGCTTATTTGAAAGAAAGATGCCAAAGTC 1200  
DB 3493 GGAATTTAGGTTTATGACAAGTAACTCAAGTGTGCTTATTTGAAAGAAAGATGCCAAAGTC 3552  
QY 1201 GAAGCAACAGAAAGAAAGGCAATATTCATTTCTTACAGTGAATGAGCAACTATGGGA 1260  
DB 3553 GAAGCAACAGAAAGAAAGGCAATATTCATTTCTTACAGTGAATGAGCAACTATGGGA 3612  
QY 1261 GCAAGTACTTCTCCATTAATAAATTAACAAATTTATATTTGGAGAAAGCCAAATGGAATCTT 1320  
DB 3613 GCAAGTACTTCTCCATTAATAAATTAACAAATTTATATTTGGAGAAAGCCAAATGGAATCTT 3672  
QY 1321 CTCAGTATCGAGAGCGGATATATTTCTGCAAAAAGTAATTCCTAATGTAACTATGGAAGA 1380  
DB 3673 CTCAGTATCGAGAGCGGATATATTTCTGCAAAAAGTAATTCCTAATGTAACTATGGAAGA 3732  
QY 1381 GAAGTAAATCGAAGGAAAGCAGATATTAATCTTCAAAATCTGAAATATCTATGATGCT 1440  
DB 3733 GAAGTAAATCGAAGGAAAGCAGATATTAATCTTCAAAATCTGAAATATCTATGATGCT 3792  
QY 1441 TCTGTTCTGTTGGAACGATGAGATATTCCTAATAAGAGCTTTTCACTATTAAGTCTCTCC 1500  
DB 3793 TCTGTTCTGTTGGAACGATGAGATATTCCTAATAAGAGCTTTTCACTATTAAGTCTCTCC 3852  
QY 1501 GAAAGAGAAATTAATCTTCCGTCAGAGATTGCTAAAGAGCAAAAGTAAATCAGAAACG 1560  
DB 3853 GAAAGAGAAATTAATCTTCCGTCAGAGATTGCTAAAGAGCAAAAGTAAATCAGAAACG 3912  
QY 1561 GATGATGTAATGTGAGAAAGTGAAGCGATTAATTCATTCGAGCTGCTGTAAGAGTGA 1620  
DB 3913 GATGATGTAATGTGAGAAAGTGAAGCGATTAATTCATTCGAGCTGCTGTAAGAGTGA 3972  
QY 1621 TTGGGGGATAGTGGTAATGGGGTTGTGGCTGCAAAATTTCTAATCTATAAAGTCTCTCC 1680  
DB 3973 TTGGGGGATAGTGGTAATGGGGTTGTGGCTGCAAAATTTCTAATCTATAAAGTCTCTCC 4032  
QY 1681 CGTATGATGTAGATGATGATCTACATCCCAAGAGCACTAAATGTGAGAGCTCATTAAC 1740  
DB 4033 CGTATGATGTAGATGATGATCTACATCCCAAGAGCACTAAATGTGAGAGCTCATTAAC 4092  
QY 1741 ATTACTAAAAATAGTGTCTGCAAAACAGAGTCTGATTTGGAACTTCCAAAGTTATGAAT 1800  
DB 4093 ATTACTAAAAATAGTGTCTGCAAAACAGAGTCTGATTTGGAACTTCCAAAGTTATGAAT 4152  
QY 1801 GATCAGCTTTATGATCAGGTCATCTAATAATCAATTTTAAATGCAATAAAACAGCGGTTT 1860  
DB 4153 GATCAGCTTTATGATCAGGTCATCTAATAATCAATTTTAAATGCAATAAAACAGCGGTTT 4212

QY 1861 GGAGGAGACGTGTCAATGAGGAATAAAGATAAGCTTACGAACTTATTTAGTCCGT 1920  
DB 4213 GGAGGAGACAGGTGTAATGAGGAATAAAGATAAGCTTATTTAGTCCGT 4272  
QY 1921 GTGTGCAACCATGAGCAATCATATATATTTCTGCTCTGTGGCAATGAGAGAGTGA 1980  
DB 4273 GTGTGCAACCATGAGCAATCATATATATTTCTGCTCTGTGGCAATGAGAGAGTGA 4332  
QY 1981 AGACTTTCTTCAGAGTGGAGGAGTAATGTAGGGCAATTAAATGAAGCTCAAAATCTT 2040  
DB 4333 AGACTTTCTTCAGAGTGGAGGAGTAATGTAGGGCAATTAAATGAAGCTCAAAATCTT 4392  
QY 2041 CGAGGACTACTGCTCAAGTGGAGGTGGCTGTACGAAAGGAAAAAGAACTTATTT 4452  
DB 4393 CGAGGACTACTGCTCAAGTGGAGGTGGCTGTACGAAAGGAAAAAGAACTTATTT 4452  
QY 2101 GGAATGCGAGGCTTTTATGGAACCTAATAAATTAATGCTTGTGCAATTCGCGAT 2160  
DB 4453 GGAATGCGAGGCTTTTATGGAACCTAATAAATTAATGCTTGTGCAATTCGCGAT 4512  
QY 2161 CATGCTGATTTGATTCGGAAGGAAAAATGTATATCAACATGAAAAATTAATGATAT 2220  
DB 4513 CATGCTGATTTGATTCGGAAGGAAAAATGTATATCAACATGAAAAATTAATGATAT 4572  
QY 2221 AAAAAATCCTTCAAAAAATGGCAAAAGTCTGTATTTGATTAATGAACTTTTAAAGAGCT 2280  
DB 4573 AAAAAATCCTTCAAAAAATGGCAAAAGTCTGTATTTGATTAATGAACTTTTAAAGAGCT 4632  
QY 2281 TTTGGAAAAAGAAACGAAATCTCAGAAATGTATCCGAAAGATTTGAATCTTGTAAAA 2340  
DB 4633 TTTGGAAAAAGAAACGAAATCTCAGAAATGTATCCGAAAGATTTGAATCTTGTAAAA 4692  
QY 2341 TTATGATGACATTTTTCAGAAAAATTTGATGAAAAACCGAGCTTTTACTTAATGCTGA 2400  
DB 4693 TTATGATGACATTTTTCAGAAAAATTTGATGAAAAACCGAGCTTTTACTTAATGCTGA 4752  
QY 2401 AGAATGACATTTATTTCTCCGATGGAACCTTCAAAAAACGAACTGTCTATAGAAATTTGA 2460  
DB 4753 AGAATGACATTTATTTCTCCGATGGAACCTTCAAAAAACGAACTGTCTATAGAAATTTGA 4812  
QY 2461 AACTATGCTCAGGAGAGAAATGAAAAATTTAGAGGAAAAATTTACCGAAAGATTTTAAAGCT 2520  
DB 4813 AACTATGCTCAGGAGAGAAATGAAAAATTTAGAGGAAAAATTTACCGAAAGATTTTAAAGCT 4872  
QY 2521 TTTTCAGAGAGATTGAGTGAATGATTAAGAACTTTGAATTTTACAGAGATGAGAAAT 2580  
DB 4873 TTTTCAGAGAGATTGAGTGAATGATTAAGAACTTTGAATTTTACAGAGATGAGAAAT 4932  
QY 2581 TATGCAAAATTTTACACTTTTACCTTCTCCGAGCTAATGAGAGAAAGAGATGTTCTTCT 2640  
DB 4933 TATGCAAAATTTTACACTTTTACCTTCTCCGAGCTAATGAGAGAAAGAGATGTTCTTCT 4992  
QY 2641 GTGGAGGAGCTGTTTCTGTGGGTGAGAACAGGAGAAATTTATAGCAAGGTATCCGTTGAAAA 2700  
DB 4993 GTGGAGGAGCTGTTTCTGTGGGTGAGAACAGGAGAAATTTATAGCAAGGTATCCGTTGAAAA 5052  
QY 2701 GGAGCTAAACTGCTGCAAAAAAGATTTTAAATATATATTAAGCTATCAATTAAGCAGAAAC 2760  
DB 5053 GGAGCTAAACTGCTGCAAAAAAGATTTTAAATATATATTAAGCTATCAATTAAGCAGAAAC 5112  
QY 2761 GTGAATTTATGTTGAAAAATTTGACTTTCGAGAGAGATCATCCGGAAGTGCAGTCGGA 2820  
DB 5113 GTGAATTTATGTTGAAAAATTTGACTTTCGAGAGAGATCATCCGGAAGTGCAGTCGGA 5172  
QY 2821 GGAAGATTAAATGTTTCAAGATGGAATAATTCAGCTATGTTAGACTTAAAGAAAGCT 2880  
DB 5173 GGAAGATTAAATGTTTCAAGATGGAATAATTCAGCTATGTTAGACTTAAAGAAAGCT 5232  
QY 2881 GAATTTATCAGAGAAAAATATTAATGAGATGATTTGAACAGACTTTTTCATGTAAGCGGA 2940  
DB 5233 GAATTTATCAGAGAAAAATATTAATGAGATGATTTGAACAGACTTTTTCATGTAAGCGGA 5292

QY 2941 TCTTTAATGCTGCTCAGTGGGAATGCAATCAATGGAATGGAAGTTATATGAGAGCT 3000  
DB 5293 TCTTTAATGCTGCTCAGTGGGAATGCAATCAATGGAATGGAAGTTATATGAGAGCT 5352  
QY 3001 ATCAGTAAAGGCAAGTTCATTGATGACGAAGCATATTTGAAAGCTTAATPAAAAATTT 3060  
DB 5353 ATCAGTAAAGGCAAGTTCATTGATGACGAAGCATATTTGAAAGCTTAATPAAAAATTT 5412  
QY 3061 GCTTTAAACATTAAGATATCTTCTGTTGGAATGCTGCGGTTTCAAGCGGGAATGCGA 3120  
DB 5413 GCTTTAAACATTAAGATATCTTCTGTTGGAATGCTGCGGTTTCAAGCGGGAATGCGA 5472  
QY 3121 ACGAAAAATGCGGCGGTGCGGAGTTGCTGTCCGTTAAATGATTAATGATTTTCAACAAA 3180  
DB 5473 ACGAAAAATGCGGCGGTGCGGAGTTGCTGTCCGTTAAATGATTAATGATTTTCAACAAA 5532  
QY 3181 GCTTCATTTGAAGATTAATGACGAAGCAAAATTAATGATTAAGATTAAGATGATGA 3240  
DB 5533 GCTTCATTTGAAGATTAATGACGAAGCAAAATTAATGATTAAGATTAAGATGATGA 5592  
QY 3241 GTAACAGTAACTGCGGGAATCTTTAGAGTGAATGCAAAAAACGACCGGAACATCAACGT 3300  
DB 5593 GTAACAGTAACTGCGGGAATCTTTAGAGTGAATGCAAAAAACGACCGGAACATCAACGT 5652  
QY 3301 ATTTCTGTTGCCGAGGAATTAATGAGTTGGAAGTAAACCGAGTGAAGAAAAACGAAA 3360  
DB 5653 ATTTCTGTTGCCGAGGAATTAATGAGTTGGAAGTAAACCGAGTGAAGAAAAACGAAA 5712  
QY 3361 TCAGAAAGAAAGCCAGAGGATTTTGTGGCAAAATCGAAAAACAAAGTGAAGCTGTAAAA 3420  
DB 5713 TCAGAAAGAAAGCCAGAGGATTTTGTGGCAAAATCGAAAAACAAAGTGAAGCTGTAAAA 5772  
QY 3421 AATAAATTTAGGATGATGATGATTCATTTACAGAAAAATTTACAAATTTACTTTCTGA 3480  
DB 5773 AATAAATTTAGGATGATGATGATTCATTTACAGAAAAATTTACAAATTTACTTTCTGA 5832  
QY 3481 GGAGTAAAAAAGCGGGGAATCTTCTCGAAGCTTTCTCATATCTCCGATTAAGAGACCG 3540  
DB 5833 GGAGTAAAAAAGCGGGGAATCTTCTCGAAGCTTTCTCATATCTCCGATTAAGAGACCG 5892  
QY 3541 TCTTTCAGTTTGGAGCTTCTGGAAGTGTTCCTTCAATTAATTTAAAAAGAAACATCT 3600  
DB 5893 TCTTTCAGTTTGGAGCTTCTGGAAGTGTTCCTTCAATTAATTTAAAAAGAAACATCT 5952  
QY 3601 GCTGTGTAATGAGTAAAGTAAATTTGAAGGAGCAATTAAGATGAGTGAAT 3660  
DB 5953 GCTGTGTAATGAGTAAAGTAAATTTGAAGGAGCAATTAAGATGAGTGAAT 6012  
QY 3661 TCTTCTGATTTCACTTTTGTGAGACATGAGGCGGATCTGTCGACTTCAATGATCAT 3720  
DB 6013 TCTTCTGATTTCACTTTTGTGAGACATGAGGCGGATCTGTCGACTTCAATGATCAT 6072  
QY 3721 ATTGGAAGTGAATATGCAACATCAATGCTGTTTATGCTGAGAGCGGCTGCTTAATTAAT 3780  
DB 6073 ATTGGAAGTGAATATGCAACATCAATGCTGTTTATGCTGAGAGCGGCTGCTTAATTAAT 6132  
QY 3781 ATTCAAAAGTAAACCAAGTCTTTGTTAAATAAGTATATTCGAATTCGCAATTAATTT 3840  
DB 6133 ATTCAAAAGTAAACCAAGTCTTTGTTAAATAAGTATATTCGAATTCGCAATTAATTT 6192  
QY 3841 AAAAGTAAATGCTTGAAGTGAAGAACTCAAGTGAAGCAGAGAGAGTGTGAAGAGCTT 3900  
DB 6193 AAAAGTAAATGCTTGAAGTGAAGAACTCAAGTGAAGCAGAGAGAGTGTGAAGAGCTT 6252  
QY 3901 AAAAGTGAAGGAGCAAGGAAAAATTAATCTTAATGGAATCTTCTGCTTATCAACTTA 3960  
DB 6253 AAAAGTGAAGGAGCAAGGAAAAATTAATCTTAATGGAATCTTCTGCTTATCAACTTA 6312  
QY 3961 GTGAACAATGAAGTTTCTGCAAAATCAAGAAATTAATACGTGACGAGAGAAATCTGAAC 4020  
DB 6313 GTGAACAATGAAGTTTCTGCAAAATCAAGAAATTAATACGTGACGAGAGAAATCTGAAC 6372  
QY 4021 CAAAAAATGGAATGTTGATGATGATCTGCTTATCAAGCGGACACCAAGTGAAGAGCTTTA 4080



QY 6241 ATCAACAAGTTTATGAGAAAACAGAGCTCTTGAGAAAATTCATATTAATGCAAAA 6300  
DB 8593 ATCAACAAGTTTATGAGAAAACAGAGCTCTTGAGAAAATTCATATTAATGCAAAA 8652  
QY 6301 CATTTATCTGTAAATCAGAGATTAACAAGATTCAATCGAGTAGAGTTCTGTGGT 6360  
DB 8653 CATTTATCTGTAAATCAGAGATTAACAAGATTCAATCGAGTAGAGTTCTGTGGT 8712  
QY 6361 GTTGTGGAATGTAGAGTAGAGCTTCTGTATCCATATTTAATAAAGAAATAC 6420  
DB 8713 GTTGTGGAATGTAGAGTAGAGCTTCTGTATCCATATTTAATAAAGAAATAC 8772  
QY 6421 AAGACAAGTTGAGAAAACATCAATGTCTGATGAAGGTTTCGAGAAAGAACTGAAAT 6480  
DB 8773 AAGACAAGTTGAGAAAACATCAATGTCTGATGAAGGTTTCGAGAAAGAACTGAAAT 8832  
QY 6481 ACAGCAGATTTCTAACAAGAAATTTCTCTTTTGGAGTCGAGTCGACAGCCGGGTA 6540  
DB 8833 ACAGCAGATTTCTAACAAGAAATTTCTCTTTTGGAGTCGAGTCGACAGCCGGGTA 8892  
QY 6541 GGAGCCGAGTGGCAGAGAACCGTTTCGTAATCAATTTGACAGAAAGACGAGATGAT 6600  
DB 8893 GGAGCCGAGTGGCAGAGAACCGTTTCGTAATCAATTTGACAGAAAGACGAGATGAT 8952  
QY 6601 GTGGAAGAGCAAAAGATTTTGGTAAAAAAGCTGAGATTAACAGCAAAAGATTAGTTCT 6660  
DB 8953 GTGGAAGAGCAAAAGATTTTGGTAAAAAAGCTGAGATTAACAGCAAAAGATTAGTTCT 9012  
QY 6661 GTTGCAATTTGAAATGCCGAGTCGAGTCGAGTCGAAAGAGCTGGAATTTGAGACGA 6720  
DB 9013 GTTGCAATTTGAAATGCCGAGTCGAGTCGAGTCGAAAGAGCTGGAATTTGAGACGA 9072  
QY 6721 GTGCGAGTTACCAAGATGAATCAACAAGAGCAAGAGTGAATAATCTAAATTTATG 6780  
DB 9073 GTGCGAGTTACCAAGATGAATCAACAAGAGCAAGAGTGAATAATCTAAATTTATG 9132  
QY 6781 ACTCGAACAAGTTAGATGTAAATAGCAAAAAATGAGATTAATCAGGTACTGGAATCGGT 6840  
DB 9133 ACTCGAACAAGTTAGATGTAAATAGCAAAAAATGAGATTAATCAGGTACTGGAATCGGT 9192  
QY 6841 TCAGCCGAGTGGAAATTCCTGACGCCGAGTATCTGAGTGGTTCTGTCAATAATATT 6900  
DB 9193 TCAGCCGAGTGGAAATTCCTGACGCCGAGTATCTGAGTGGTTCTGTCAATAATATT 9252  
QY 6901 GCAAAATAGTAGAAAACAGATATCGATCACTACTTTCACTGATGTAAAT 6960  
DB 9253 GCAAAATAGTAGAAAACAGATATCGATCACTACTTTCACTGATGTAAAT 9312  
QY 6961 GTAAAAAGTCTTATATAAATTTTCGAATTCCTTGAACAGCCGGTGGAGAGCCGAGGTCTT 7020  
DB 9313 GTAAAAAGTCTTATATAAATTTTCGAATTCCTTGAACAGCCGGTGGAGAGCCGAGGTCTT 9372  
QY 7021 GGAGCAGTTACCGGAGTGGTTCTGTTAACATATAATAGTTCTGTAGTACGTGAGTT 7080  
DB 9373 GGAGCAGTTACCGGAGTGGTTCTGTTAACATATAATAGTTCTGTAGTACGTGAGTT 9432  
QY 7081 CACAATACTGTGATTTGACTTCCTGACGAGAAAAGTAAATGTAACGGCAAAAAGAGAA 7140  
DB 9433 CACAATACTGTGATTTGACTTCCTGACGAGAAAAGTAAATGTAACGGCAAAAAGAGAA 9492  
QY 7141 AAAAATATTAAGCAAAACAGACGAAATGACGAATGGAGAGACGACCAATCGAGCAAT 7200  
DB 9493 AAAAATATTAAGCAAAACAGACGAAATGACGAATGGAGAGACGACCAATCGAGCAAT 9552  
QY 7201 GTCTTGATAATTAATTTTGGAAACAGCTGTAGAAAGATGAAGAAAATTTCTGAAGAAAAGGA 7260  
DB 9553 GTCTTGATAATTAATTTTGGAAACAGCTGTAGAAAGATGAAGAAAATTTCTGAAGAAAAGGA 9612  
QY 7261 ACAGAAGTTTAAAACTTTAGACGAATTAACAAAGAACAGATTAATAAAGTAAATGAT 7320  
DB 9613 ACAGAAGTTTAAAACTTTAGACGAATTAACAAAGAACAGATTAATAAAGTAAATGAT 9672

QY 7321 GCTAGAAAAAATCTTACAATCAGCAGATATTTCTACAGAGATATCTTGTAAAAAGCG 7380  
DB 9673 GCTAGAAAAAATCTTACAATCAGCAGATATTTCTACAGAGATATCTTGTAAAAAGCG 9732  
QY 7381 GATAGAGAGATATCTCAGGAGAGAGAAATTAAGCCATTGTGAAGACTTCTGATATTAAT 7440  
DB 9733 GATAGAGAGATATCTCAGGAGAGAGAAATTAAGCCATTGTGAAGACTTCTGATATTAAT 9792  
QY 7441 GAAAAAATGTATGATTAACAACAGAGACAAAGATTAATATCATCTTACGTGGTGGT 7500  
DB 9793 GAAAAAATGTATGATTAACAACAGAGACAAAGATTAATATCATCTTACGTGGTGGT 9852  
QY 7501 GGAACCTGCAGGCTTGCTTCGATCAGGAACAGTGGCAGTTACAAATTTAAAAAGAAAT 7560  
DB 9853 GGAACCTGCAGGCTTGCTTCGATCAGGAACAGTGGCAGTTACAAATTTAAAAAGAAAT 9912  
QY 7561 TCCGGAGTTACTGTTGAAAAATCTTTTGTGAAGCAGTGAATAAGTAAATGTTAGATCG 7620  
DB 9913 TCCGGAGTTACTGTTGAAAAATCTTTTGTGAAGCAGTGAATAAGTAAATGTTAGATCG 9972  
QY 7621 GATATTAACAGAAATGTTGCTTTAACAGATATCAAGGTCCTGTAGAGCATTTGGAAAT 7680  
DB 9973 GATATTAACAGAAATGTTGCTTTAACAGATATCAAGGTCCTGTAGAGCATTTGGAAAT 10032  
QY 7681 GGAGCTGCCTATGACAGAAATTAATCTAATGGAAGATCAAAATATCATATTAATAATCT 7740  
DB 10033 GGAGCTGCCTATGACAGAAATTAATCTAATGGAAGATCAAAATATCATATTAATAATCT 10092  
QY 7741 AAGCTATTAAGAAAAAATATTTGATGTATTTGTAAAAAGATTAATTCGAAATTTGAGACGGA 7800  
DB 10093 AAGCTATTAAGAAAAAATATTTGATGTATTTGTAAAAAGATTAATTCGAAATTTGAGACGGA 10152  
QY 7801 GCAAAAAGATTAACCGTAGAGCCGTAGTGTCCGAGCCATTAATCTCAAAACCAAAAGAT 7860  
DB 10153 GCAAAAAGATTAACCGTAGAGCCGTAGTGTCCGAGCCATTAATCTCAAAACCAAAAGAT 10212  
QY 7861 GAAATGAATTCAGAGTTGAAATTTGAGAGATATTTTCAATGAAGAAATAGATGACT 7920  
DB 10213 GAAATGAATTCAGAGTTGAAATTTGAGAGATATTTTCAATGAAGAAATAGATGACT 10272  
QY 7921 AGCCCTTCTAAAGAAATTTGAGAGAAATCAATGTCAAAAGTGAATAAAGAAAAACAGAGTG 7980  
DB 10273 AGCCCTTCTAAAGAAATTTGAGAGAAATCAATGTCAAAAGTGAATAAAGAAAAACAGAGTG 10332  
QY 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGTAGAGGGGACAGAAATTAATTTCCGA 8040  
DB 10333 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGTAGAGGGGACAGAAATTAATTTCCGA 10392  
QY 8041 GCAAAAAGATGCGGAAAGCTCTTAATTTGAAGTTAGTACAAAATCCGGAAGAAATTTTTT 8100  
DB 10393 GCAAAAAGATGCGGAAAGCTCTTAATTTGAAGTTAGTACAAAATCCGGAAGAAATTTTTT 10452  
QY 8101 CATGCAATATGTGATTAATGGAAGCAACATTAATAATGAAGTAAACAGCAGTTCTAAA 8160  
DB 10453 CATGCAATATGTGATTAATGGAAGCAACATTAATAATGAAGTAAACAGCAGTTCTAAA 10512  
QY 8161 GCAAGTAACAGGTTCTGTATTTGAGAGAGTTGAGATCAACAGGACAGACTACTGCTGCA 8220  
DB 10513 GCAAGTAACAGGTTCTGTATTTGAGAGAGTTGAGATCAACAGGACAGACTACTGCTGCA 10572  
QY 8221 GGTAAAACTAATGTGAAAGTTGAGAGAGAAATTTGTTCAAGCAAAATGATGAAATGCA 8280  
DB 10573 GGTAAAACTAATGTGAAAGTTGAGAGAGAAATTTGTTCAAGCAAAATGATGAAATGCA 10632  
QY 8281 ATTTCTAAAGTAGAAGGTTTGGATGAGAAATTAAGTAACTGCTAAATCTTCTGTAGATCA 8340  
DB 10633 ATTTCTAAAGTAGAAGGTTTGGATGAGAAATTAAGTAACTGCTAAATCTTCTGTAGATCA 10692  
QY 8341 GGAATATGAGAGAGAAATTCGAGAGCAGAGATGAAATCTTCTACAGACAAAGTAAATACT 8400  
DB 10693 GGAATATGAGAGAGAAATTCGAGAGCAGAGATGAAATCTTCTACAGACAAAGTAAATACT 10752  
QY 8401 GAATCCGTAGTTGCTTTAGAGAAAGCAAGATTTAGAAATTAATGATTAACAAAAAATAT 8460

```
Db 10753 GAATCCGATGTCGTTTACGAAGCAAGATATGAAAAATTAATTACACAAAAAATAT 10812
Qy 8461 ATTTGAGAGTCAATGCTCTTGGCTTTAAATGATACAAAGAAAGAACCAATTAATAC 8520
Db 10813 ATTTGAGAGTCAATGCTCTTGGCTTTAAATGATACAAAGAAAGAACCAATTAATAC 10872
Qy 8521 TTAGCGGTAGCCGATGTCATGACACAGAAACAAACCAATTCAGATCAACAAG 8580
Db 10873 TTAGCGGTAGCCGATGTCATGACACAGAAACAAACCAATTCAGATCAACAAG 10932
Qy 8581 TTAACCTTACACCTGTAAATGAGAGAAACGATCTCAACCTTCGCAAAAGCTTTGGCT 8640
Db 10933 TTAACCTTACACCTGTAAATGAGAGAAACGATCTCAACCTTCGCAAAAGCTTTGGCT 10992
Qy 8641 AAAAAATGAAAAATTAAGAAATGTAAGAACTGAGAGACCTTAGTCGAGCGGAAACA 8700
Db 10993 AAAAAATGAAAAATTAAGAAATGTAAGAACTGAGAGACCTTAGTCGAGCGGAAACA 11052
Qy 8701 GCAGCCGTTGAAAAATTTATACAAAGATCTACAGAGCACTTGGTTCAGAAATTTGGGA 8760
Db 11053 GCAGCCGTTGAAAAATTTATACAAAGATCTACAGAGCACTTGGTTCAGAAATTTGGGA 11112
Qy 8761 ATTGAGATTAATTAAGAAACGATTCAGAGATTAATACGATTTGAAGTCAACGAGAC 8820
Db 11113 ATTGAGATTAATTAAGAAACGATTCAGAGATTAATACGATTTGAAGTCAACGAGAC 11172
Qy 8821 GGAACCAAGAGAGCTTGTCTGCAAGAAAGATGTTATTTCTGTGAAAAATTAATTCAGAG 8880
Db 11173 GGAACCAAGAGAGCTTGTCTGCAAGAAAGATGTTATTTCTGTGAAAAATTAATTCAGAG 11232
Qy 8881 GAAACAAAATCATCATTTGAAGATTAAGCCAGAAATTTGGAACCGGAAAGTTAATGTA 8940
Db 11233 GAAACAAAATCATCATTTGAAGATTAAGCCAGAAATTTGGAACCGGAAAGTTAATGTA 11292
Qy 8941 GATGCTTTGAATGAACCTGATGATGATCTACAAAGAAAAAGTGTGCTATGTTGTAAT 9000
Db 11293 GATGCTTTGAATGAACCTGATGATGATCTACAAAGAAAAAGTGTGCTATGTTGTAAT 11352
Qy 9001 GGTATTTGAAAAATGTTGATGATTAATGATTAAGAAAAATGTAAGAACCAATTCGGA 9060
Db 11353 GGTATTTGAAAAATGTTGATGATTAATGATTAAGAAAAATGTAAGAACCAATTCGGA 11412
Qy 9061 AGACATGCTATGTAAGAACTACTGAGAAAAACAAGATATCAAGCACTTACAGAGCAAAA 9120
Db 11413 AGACATGCTATGTAAGAACTACTGAGAAAAACAAGATATCAAGCACTTACAGAGCAAAA 11472
Qy 9121 GTAAATATTTCTGAAAAAGAGACGCTGCACTGCAATTCGAATGTACACAT 9180
Db 11473 GTAAATATTTCTGAAAAAGAGACGCTGCACTGCAATTCGAATGTACACAT 11532
Qy 9181 TCCCAATGAGATGATTAATTAATTTGGCAAGACGATGATCTTCTCAATTAATAC 9240
Db 11533 TCCCAATGAGATGATTAATTAATTTGGCAAGACGATGATCTTCTCAATTAATAC 11592
Qy 9241 AAAAAATTCAAAAAATTAATTAATCTTAGCATCAAGATGAATCGAATGTGAATGTTGAT 9300
Db 11593 AAAAAATTCAAAAAATTAATTAATCTTAGCATCAAGATGAATCGAATGTGAATGTTGAT 11652
Qy 9301 GGGGCTGCTGAAGCAAGAGTGCAGAGCCAAAGCGACATTAAGTGAAGATCAAAATA 9360
Db 11653 GGGGCTGCTGAAGCAAGAGTGCAGAGCCAAAGCGACATTAAGTGAAGATCAAAATA 11712
Qy 9361 AATAAGACTAATTAATGTTGATTTAGCAGAAAAATTAAGACAGAGAAACATCATGTA 9420
Db 11713 AATAAGACTAATTAATGTTGATTTAGCAGAAAAATTAAGACAGAGAAACATCATGTA 11772
Qy 9421 TATGCGGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480
Db 11773 TATGCGGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 11832
Qy 9481 GCCAAAGCTCATGCTGACAGCTCTTCGCAACTGCACTATTGAAAAAATGAAGTAAATA 9540
```

```
Db 11833 GCCAAAGCTCATGCTGACAGCTCTTCGCAACTGCACTATTGAAAAAATGAAGTAAATA 11892
Qy 9541 TTTAATATATGAGATCCGGAATTTAAAAATATATGCAAGATTGGAAGGAAAGCTAAT 9600
Db 11893 TTTAATATATGAGATCCGGAATTTAAAAATATATGCAAGATTGGAAGGAAAGCTAAT 11952
Qy 9601 AAAAAACGTGAGATGATCTAATCAGGTAGACTGATTAACGATTAATTAATTAATTAAT 9660
Db 11953 AAAAAACGTGAGATGATCTAATCAGGTAGACTGATTAACGATTAATTAATTAATTAAT 12012
Qy 9661 TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATTAAGAGAGGAGAAAAAGGAAA 9720
Db 12013 TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATCAAGAGAGGAGAAAAAGGAAA 12072
Qy 9721 AAATGA 9726
Db 12073 AAATGA 12078

RESULT 4
AR442022
LOCUS AR442022 2780 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 10 from patent US 6669940.
ACCESSION AR442022
VERSION AR442022.1 GI:42669173
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2780)
AUTHORS Negara,J.A., T.G., Stewart,G.C., Narayanan,S.K. and Chengappa,M.M.
TITLE Recombinant fusoobacterium necrophorum leukotoxin vaccine and
JOURNAL preparation thereof
Patent: US 6669940-A 10-30-DEC-2003;
Kansas University Research Foundation; Manhattan, KS
FEATURES
location/Qualifiers
source 1..2780
/mol_type="genomic DNA"

ORIGIN
Query Match 28.6%; Score 2780; DB 6; Length 2780;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

919 GGGAGATTTATTAAGAAAGAGGGGAAAAAGAACTTATACACTCTTTAAGTTATCA 978
Db 1 GGGAGATTTATTAAGAAAGAGGGGAAAAAGAACTTATACACTCTTTAAGTTATCA 60
Qy 979 GATGTGGAAGCTCCGTAAAGTAAATAAGAAAGATCATAGAAAGATGTTGACAT 1038
Db 61 GATGTGGAAGCTCCGTAAAGTAAATAAGAAAGATCATAGAAAGATGTTGACAT 120
Qy 1039 ACAGCTGAAGCAAGAAATTTCTATGATGCAACTTTAGTTACTAAGCTTGCACACCTCT 1098
Db 121 ACAGCTGAAGCAAGAAATTTCTATGATGCAACTTTAGTTACTAAGCTTGCACACCTCT 180
Qy 1099 TTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAATGATTTTGTAGTTATG 1158
Db 181 TTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAATGATTTTGTAGTTATG 240
Qy 1159 ACAAGTAACTCAGAGTGTGTTATTGGAAGATGTCGAAGTCGAAGCAACGAAGAAAG 1218
Db 241 ACAAGTAACTCAGAGTGTGTTATTGGAAGATGTCGAAGTCGAAGCAACGAAGAAAG 300
Qy 1219 GCAATATTCATTTCTTACAGTGAAGTAAAGCAACTATAGGAGCAGCTACTTCTCATTA 1278
Db 301 GCAATATTCATTTCTTACAGTGAAGTAAAGCAACTATAGGAGCAGCTACTTCTCATTA 360
Qy 1279 AAAATTTACCAATTTATATTTGAGAGAACCAATGGAACCTTCTAGATATGAGCGGGA 1338
Db 361 AAAATTTACCAATTTATATTTGAGAGAACCAATGGAACCTTCTAGATATGAGCGGGA 420
```

QY 1339 TATATTTTCGCAAAAAGTAATTCGAATGTAACTATTTGAAGAGAGATAAATAATCGAAGGA 1398  
Db 421 TATATTTTCGCAAAAAGTAATTCGAATGTAACTATTTGAAGAGAGATAAATAATCGAAGGA 480  
QY 1399 AGAGCAGATATTACTCTCAAAATCTGAAAAATACATATGTAGTCTTCTGTTCTGTTGGAACG 1458  
Db 481 AGAGCAGATATTACTCTCAAAATCTGAAAAATACATATGTAGTCTTCTGTTCTGTTGGAACG 540  
QY 1459 ATGAGAGATTCGAATTAAGTAGTCTTTCAGATTTGTGTGACGGAAGAGAAAAATTAATCT 1518  
Db 541 ATGAGAGATTCGAATTAAGTAGTCTTTCAGATTTGTGTGACGGAAGAGAAAAATTAATCT 600  
QY 1519 TCCCGCAAGATTGCTTAAGSACAAAAGTAGAATCAGAAAACGGAGAGTGAATGTGAGA 1578  
Db 601 TCCCGCAAGATTGCTTAAGSACAAAAGTAGAATCAGAAAACGGAGAGTGAATGTGAGA 660  
QY 1579 AGTGAAGCGATTAAATTCCTGAGCTGCTGTAAGAGTGAATGGGGGATAGTGTAT 1638  
Db 661 AGTGAAGCGATTAAATTCCTGAGCTGCTGTAAGAGTGAATGGGGGATAGTGTAT 720  
QY 1639 GGGGTTGTGCTGCAAAATATTTCTAACTATTAATGCTTCTCCGTAATAGATGTAGTGA 1698  
Db 721 GGGGTTGTGCTGCAAAATATTTCTAACTATTAATGCTTCTCCGTAATAGATGTAGTGA 780  
QY 1699 TATCTACATGCCAAGAGCGACTTAAATGTGAGGCTCAACATTAATCTAAAAATAGTCT 1758  
Db 781 TATCTACATGCCAAGAGCGACTTAAATGTGAGGCTCAACATTAATCTAAAAATAGTCT 840  
QY 1759 CTGCAAAACAGATCTGATTTGGGAACCTTCAAGTTTATGATGATCACTTTATGATCA 1818  
Db 841 CTGCAAAACAGATCTGATTTGGGAACCTTCAAGTTTATGATGATCACTTTATGATCA 900  
QY 1819 GGTCAATCTAAAAATCAATTTTAAATGCAATTAACAGCGGTTTGGAGAGACAGTGTCAAT 1878  
Db 901 GGTCAATCTAAAAATCAATTTTAAATGCAATTAACAGCGGTTTGGAGAGACAGTGTCAAT 960  
QY 1879 GAGGAATTAAGATTAAGCTAACGAATTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1938  
Db 961 GAGGAATTAAGATTAAGCTAACGAATTTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1939 AATCATTAATTAATTCGCTTCTGTCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1998  
Db 1021 AATCATTAATTAATTCGCTTCTGTCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
QY 1999 GAAAGGATTAATGTAAGGCAATTAATGAAGCTCAAAATCTTGCAGCGACTAGCTCAAGT 2058  
Db 1081 GAAAGGATTAATGTAAGGCAATTAATGAAGCTCAAAATCTTGCAGCGACTAGCTCAAGT 1140  
QY 2059 GGAAGTGTGCTGTACGAAAGGAAGAAAAAGAACTTATTTGGAATTCAGCAGATTTT 2118  
Db 1141 GGAAGTGTGCTGTACGAAAGGAAGAAAAAGAACTTATTTGGAATTCAGCAGATTTT 1200  
QY 2119 TATGGAATCTATTAATTAATGCTTCTGTCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2178  
Db 1201 TATGGAATCTATTAATTAATGCTTCTGTCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
QY 2179 GAAAGAAAAATTTGATATCAACAGTGAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 2238  
Db 1261 GAAAGAAAAATTTGATATCAACAGTGAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320  
QY 2239 GCAAGTCTGTTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2298  
Db 1321 GCAAGTCTGTTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380  
QY 2299 ACTCCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2358  
Db 1381 ACTCCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440  
QY 2359 GAAAAATTTGATGAGAAAAACCGAGCTTTTAACTAAATGTTGAAAGAAATGAACAATTAATCTT 2418  
Db 1441 GAAAAATTTGATGAGAAAAACCGAGCTTTTAACTAAATGTTGAAAGAAATGAACAATTAATCTT 1500  
QY 2419 CCGGATGGAACCTTCAAAAACAGGAACCTGTAATGAATAATTCGAACCTATGTTTCAGGAGAA 2478

Db 1501 CCGATGGAACCTTCAAAAACAGGAACCTGTAATGAATAATTCGAACCTATGTTTCAGGAGAA 1560  
QY 2479 ATGAAAAATTTAGAGAAAAATTTACCGAAGAGATTTAAAGCTTTTCAGAAAGATTTGAGT 2538  
Db 1561 ATGAAAAATTTAGAGAAAAATTTACCGAAGAGATTTAAAGCTTTTCAGAAAGATTTGAGT 1620  
QY 2539 GGAAGTAAAGAACTTTGAATTTTACAGAGTAGAGAAATTAATCAAAATTTTCACACT 2598  
Db 1621 GGAAGTAAAGAACTTTGAATTTTACAGAGTAGAGAAATTTACAAATTTTTCACACT 1680  
QY 2599 TTTACCTTCCGAGCTTAATGAGAAAAAGATGTTTCTTCTGTCGAGAGAGCTTTTCG 2658  
Db 1681 TTTACCTTCCGAGCTTAATGAGAAAAAGATGTTTCTTCTGTCGAGAGAGCTTTTCG 1740  
QY 2659 TGGGTAGAACAGAGAAATTAATGCAAGGTATTCGTTGAAAAAGAGCTTAACTTCTGCA 2718  
Db 1741 TGGGTAGAACAGAGAAATTAATGCAAGGTATTCGTTGAAAAAGAGCTTAACTTCTGCA 1800  
QY 2719 AAAAAAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2778  
Db 1801 AAAAAAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1860  
QY 2779 ATTGACCTTGCGAAGCAGTACATCCGGAAGTGCAGTCCGAGAAAGATTAATGTTCA 2838  
Db 1861 ATTGACCTTGCGAAGCAGTACATCCGGAAGTGCAGTCCGAGAAAGATTAATGTTCA 1920  
QY 2839 AGATCAAAAATTAATGCTATGCTAGTACAGCTTAAGAAAAAGCTGAATTAATCAGAGAAAT 2898  
Db 1921 AGATCAAAAATTAATGCTATGCTAGTACAGCTTAAGAAAAAGCTGAATTAATCAGAGAAAT 1980  
QY 2899 ATTAATGCAATGCTTGAACAGACTTTTTCATGTAAGCGGATCTTTTAAATGCTGCTCA 2958  
Db 1981 ATTAATGCAATGCTTGAACAGACTTTTTCATGTAAGCGGATCTTTTAAATGCTGCTCA 2040  
QY 2959 GGTGGGAATGCAATTAATGGAATGGGAAGTTTAAAGAGATGATCAGTAAGCAAGATT 3018  
Db 2041 GGTGGGAATGCAATTAATGGAATGGGAAGTTTAAAGAGATGATCAGTAAGCAAGATT 2100  
QY 3019 TCCATTTGATGAGCAAGATTTTGAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3078  
Db 2101 TCCATTTGATGAGCAAGATTTTGAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160  
QY 3079 GATACCTTCTGTTGGAATGCTCCGCTTCAAGCGGGAATGGAACGAAAAATCCGCGGCTC 3138  
Db 2161 GATACCTTCTGTTGGAATGCTCCGCTTCAAGCGGGAATGGAACGAAAAATCCGCGGCTC 2220  
QY 3139 GGGGTTGCTGTTGCGGTAATGATTAATGATTAATTTCAACAAAGCTTCCATTGAAGATTAAT 3198  
Db 2221 GGGGTTGCTGTTGCGGTAATGATTAATGATTAATTTCAACAAAGCTTCCATTGAAGATTAAT 2280  
QY 3199 GACGAAGCAAAAGTAATATGATTAAGATTAAGATGATGAAGTAACAGTAACTGCGGAA 3258  
Db 2281 GACGAAGCAAAAGTAATATGATTAAGATTAAGATGATGAAGTAACAGTAACTGCGGAA 2340  
QY 3259 TCTTTAGAAATGATGCAAAAACGACCGGAACATCAACAGTATTTCTGTTCCGAGGA 3318  
Db 2341 TCTTTAGAAATGATGCAAAAACGACCGGAACATCAACAGTATTTCTGTTCCGAGGA 2400  
QY 3319 ATTAATTAAGGTTGGAAGTAAACGAGTGAAGAAAAACGAAATTCAGAAAGAACAGAG 3378  
Db 2401 ATTAATTAAGGTTGGAAGTAAACGAGTGAAGAAAAACGAAATTCAGAAAGAACAGAG 2460  
QY 3379 GGAATTTTTCGCAAAATCGGAACAAAGTGAACCTGTAAATAATTAATTAATTAATTAATTAAT 3438  
Db 2461 GGAATTTTTCGCAAAATCGGAACAAAGTGAACCTGTAAATAATTAATTAATTAATTAATTAAT 2520  
QY 3439 ATGATTTCAATTAACGAAAAATTAACAAATTTCAATTTCTGAAGAGATTAATTAATTAATTAAT 3498  
Db 2521 ATGATTTCAATTAACGAAAAATTAACAAATTTCAATTTCTGAAGAGATTAATTAATTAATTAAT 2580  
QY 3499 AATCTTCCCTTGAAGCTTTCTGATCTCCGGAATTAAGAGACGCTTTTCAGTTTGGAGCT 3558

Db 2581 AATCTTCCTTGAAGCTTTCATACCTCCGATTAAGAACCCTCTTTCAGTTGGAGCT 2640  
Qy 3559 TCTGGAAGTGTCTTTCATATATTAATAAGAAACATCTGCTGTGTAGATGAGTA 3618  
Db 2641 TCTGGAAGTGTCTTTCATATATTAATAAGAAACATCTGCTGTGTAGATGAGTA 2700  
Qy 3619 AAGATTAATTTGAAGGAGCAATAATAAGTAGAGTACTTCTTCTGATTTCACTTT 3678  
Db 2701 AAGATTAATTTGAAGGAGCAATAATAAGTAGAGTACTTCTTCTGATTTCACTTT 2760  
Qy 3679 GTTGGAGCATGGGGCGGATC 3698  
Db 2761 GTTGGAGCATGGGGCGGATC 2780

RESULT 5  
AR442025  
LOCUS AR442025 2322 bp DNA linear PAT 20-FEB-2004  
DEFINITION Sequence 13 from patent US 6669940.  
ACCESSION AR442025  
VERSION AR442025.1 GI:42669176  
KEYWORDS  
SOURCE  
ORGANISM  
Unclassified.

REFERENCE  
1 (bases 1 to 2322)  
Nagaraja, T.G., Stewart, G.C., Narayanan, S.K. and Chengappa, M.M.  
TITLE Recombinant fimbriated bacterium necrophorum leukotoxin vaccine and  
Preparation thereof  
Patent: US 6669940-A 13 30-DEC-2003;  
JOURNAL Kansas University Research Foundation, Manhattan, KS  
FEATURES  
source  
1..2322  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 23.9%; Score 2322; DB 6; Length 2322;  
Best Local Similarity 100.0%; Pred. No. 4e-296;  
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7405 GGAATTTAAAGCCATTGTGAAGCTTCTGATATTATTTGAAAAAATGTAGATTTTACAACA 7464  
Db 1 GGAATTTAAAGCCATTGTGAAGCTTCTGATATTATTTGAAAAAATGTAGATTTTACAACA 60  
Qy 7465 GAGGACAAGATTAATCACTTCTAGTGTGGAGTGGCACTGCAAGTCTTGCTCCGCA 7524  
Db 61 GAGGACAAGATTAATCACTTCTAGTGTGGAGTGGCACTGCAAGTCTTGCTCCGCA 120  
Qy 7525 TCAGGACAAGTGGCACTTCAATAATTTAAAGAAATTCGAGATTACTGTGAAATTTCT 7584  
Db 121 TCAGGACAAGTGGCACTTCAATAATTTAAAGAAATTCGAGATTACTGTGAAATTTCT 180  
Qy 7585 TTTGTGAAAAGCAGCTGAAAAAGTAAATGTTAGATCGATATTACAGAAATGTTGCTTTA 7644  
Db 181 TTTGTGAAAAGCAGCTGAAAAAGTAAATGTTAGATCGATATTACAGAAATGTTGCTTTA 240  
Qy 7645 ACAGCATATCAAGTCTCTAGAGCAATGGGAATTGAGCTGCTATGCAAGTTAAT 7704  
Db 241 ACAGCATATCAAGTCTCTAGAGCAATGGGAATTGAGCTGCTATGCAAGTTAAT 300  
Qy 7705 TCTAAATGGAAGTCAATATCAATATTAATAATTTCTAAGCTATTAGAAAAAATATTGAT 7764  
Db 301 TCTAAATGGAAGTCAATATCAATATTAATAATTTCTAAGCTATTAGAAAAAATATTGAT 360  
Qy 7765 GTTATTTGTAAGATTAATCGAATTGAGCGGAAACAAAGATTAACCGTAGAGCG 7824  
Db 361 GTTATTTGTAAGATTAATCGAATTGAGCGGAAACAAAGATTAACCGTAGAGCG 420  
Qy 7825 GTTAGCTGCCGAGCACTTATCTCAAAAGCAAGAAATGAATGAATTGAGAGTTGAAT 7884  
Db 421 GTTAGCTGCCGAGCACTTATCTCAAAAGCAAGAAATGAATGAATTGAGAGTTGAAT 480

Qy 7885 GAGAGAGTATTTTCAATGAAGAAATAGAGTAATAGACCTTCTTAAGAAATTTGAAGA 7944  
Db 481 GAGAGAGTATTTTCAATGAAGAAATAGAGTAATAGACCTTCTTAAGAAATTTGAAGA 540  
Qy 7945 GAAATCAATGTCAAGTGGAAAAAGAAACAGAGTACTGCTGAATCTCAAGAGCTTCT 8004  
Db 541 GAAATCAATGTCAAGTGGAAAAAGAAACAGAGTACTGCTGAATCTCAAGAGCTTCT 600  
Qy 8005 GTAGAGCAGTAGCAGGAGGACAGAAATTTATTTCCGAAAGCAAAAGATGCCGAAAGCTTAT 8064  
Db 601 GTAGAGCAGTAGCAGGAGGACAGAAATTTATTTCCGAAAGCAAAAGATGCCGAAAGCTTAT 660  
Qy 8065 TTGAAAGTTAGTACAAATCCGAAAGAGTATTTTCTATGCAAGATTAATGTGAATTTGAA 8124  
Db 661 TTGAAAGTTAGTACAAATCCGAAAGAGTATTTTCTATGCAAGATTAATGTGAATTTGAA 720  
Qy 8125 GCAACACATTAATAAGAAAGTAAACAGCACTTCTTAAGCAATGACAGGTTCTGATTTGGGA 8184  
Db 721 GCAACACATTAATAAGAAAGTAAACAGCACTTCTTAAGCAATGACAGGTTCTGATTTGGGA 780  
Qy 8185 GAGGTTGAGTCAACCAAGCAGAAAGCTAGCTGCTGCAAGTAAACCTATGATGAAGTTGAG 8244  
Db 781 GAGGTTGAGTCAACCAAGCAGAAAGCTAGCTGCTGCAAGTAAACCTATGATGAAGTTGAG 840  
Qy 8245 GAAAGAAATTTTGTTCAGAACAAATGATTAATGCAATTTCTTAAGTAAAGTTTGAT 8304  
Db 841 GAAAGAAATTTTGTTCAGAACAAATGATTAATGCAATTTCTTAAGTAAAGTTTGAT 900  
Qy 8305 GAAATTAAGTAACTGCTTAATCTTCTGTAGTATTAAGAAATGAGAGAGAAATTTGCCGGA 8364  
Db 901 GAAATTAAGTAACTGCTTAATCTTCTGTAGTATTAAGAAATGAGAGAGAAATTTGCCGGA 960  
Qy 8365 GCAGAGTGAATTAATCTTCAACAGCAAGATTAATCTGAATCCGATTCGTTTACGAAG 8424  
Db 961 GCAGAGTGAATTAATCTTCAACAGCAAGATTAATCTGAATCCGATTCGTTTACGAAG 1020  
Qy 8425 CAAGATTAAGAAATTAATGATTAACAACAAAAATTAATTTAGAAATGCAATGCTTGTCT 8484  
Db 1021 CAAGATTAAGAAATTAATGATTAACAACAAAAATTAATTTAGAAATGCAATGCTTGTCT 1080  
Qy 8485 TTTAAATGATCAAGAAATGAAGCCAAATTAAGATTTTAAAGCGGTAGCCGCTGTGATGCA 8544  
Db 1081 TTTAAATGATCAAGAAATGAAGCCAAATTAAGATTTTAAAGCGGTAGCCGCTGTGATGCA 1140  
Qy 8545 CAAGGACAAACAAAGCACTTTACGAGTCAACAAAGTTAATCTTCAACATGTAATGGA 8604  
Db 1141 CAAGGACAAACAAAGCACTTTACGAGTCAACAAAGTTAATCTTCAACATGTAATGGA 1200  
Qy 8605 GAAAACGATCTCAAATTCGTGCAAAAGCTTTGCTTAAATTAAGAAATTAATGAAATGTA 8664  
Db 1201 GAAAACGATCTCAAATTCGTGCAAAAGCTTTGCTTAAATTAAGAAATTAATGAAATGTA 1260  
Qy 8665 AAGGAATCTGAGAGAGCTTATAGTGGAGCGGAAACAGAGCCGTTGAAATTTATCAAG 8724  
Db 1261 AAGGAATCTGAGAGAGCTTATAGTGGAGCGGAAACAGAGCCGTTGAAATTTATCAAG 1320  
Qy 8725 AGTACTACAGAGAGCTGTTGTCAGAGAAATTTGAGAAATTTGAGATTAATTTGAAACGAT 8784  
Db 1321 AGTACTACAGAGAGCTTGTGTCAGAGAAATTTGAGAAATTTGAGATTAATTTGAAACGAT 1380  
Qy 8785 GCAAGAGATTAATGCAATTTGTAAGATCAACGAGAGCGAAACCAAGAGAGTCTTGTCCGA 8844  
Db 1381 GCAAGAGATTAATGCAATTTGTAAGATCAACGAGAGCGAAACCAAGAGAGTCTTGTCCGA 1440  
Qy 8845 AAGAAATGATTTTCTGTGAAAAATTAATTTTCAAGGAGAAACAAATTCATCTGAAAGAT 8904  
Db 1441 AAGAAATGATTTTCTGTGAAAAATTAATTTTCAAGGAGAAACAAATTCATCTGAAAGAT 1500  
Qy 8905 AAAAGCAGAAATTTGGAACCGGAAAGTAAATGTAATGCTTGAATGAACCTTGAATGTA 8964  
Db 1501 AAAAGCAGAAATTTGGAACCGGAAAGTAAATGTAATGCTTGAATGAACCTTGAATGTA 1560  
Qy 8965 GATCTACAGAGAAAGAGTGTGCTATGTGGAATTTGTAATTTGAAATGTTGAATGTAAT 9024



Qy	4573	ACGGAATTAATCCGCTGGAGAGTACGTAATGCAATCACTGTTAAAAATTAATTTAA	4632
Db	1021	ACGGAATTAATCCGCTGGAGAGTACGTAATGCAATCACTGTTAAAAATTTTAA	1080
Qy	4633	GCAGAAATGAGTGAAGCAATTAAGGAACCGGAGAGGATAAATTCATCCGAAACATGTA	4692
Db	1081	GCAGAAATGAGTGAAGCAATTAAGGAACCGGAGAGGATAAATTCATCCGAAACATGTA	1140
Qy	4693	AATGTGAGAGCAAAATCATCTACTGTTGTGTGAATGCGGCTTGTGACCTTGCTATCAGC	4752
Db	1141	AATGTGAGAGCAAAATCATCTACTGTTGTGTGAATGCGGCTTGTGACCTTGCTATCAGC	1200
Qy	4753	AAAGATGCTTTTTCAGAGAAATGGAATCTGGAGCATGGCAAGACTTATCAAAATGACAGAT	4812
Db	1201	AAAGATGCTTTTTCAGAGAAATGGAATCTGGAGCATGGCAAGACTTATCAAAATGACAGAT	1260
Qy	4813	GCAAAAGTGGATTAAGGAAGAAATTTCTGTGATTCCTTAAATGTGAACCGCAATTAATTC	4872
Db	1261	GCAAAAGTGGATTAAGGAAGAAATTTCTGTGATTCCTTAAATGTGAACCGCAATTAATTC	1320
Qy	4873	ATTCTTGGGGTGAATGTTGCGGGAAACCATTTGCCGGTCTCTTCTACGGCGGTAGAGCT	4932
Db	1321	ATTCTTGGGGTGAATGTTGCGGGAAACCATTTGCCGGTCTCTTCTACGGCGGTAGAGCT	1380
Qy	4933	GCTTTTGGCAATTAATCTCTTCAATAAATTAATTAATTAATTAATTAATTAATTAATTAAT	4992
Db	1381	GCTTTTGGCAATTAATCTCTTCAATAAATTAATTAATTAATTAATTAATTAATTAATTAAT	1440
Qy	4993	AATCCTTTAGTGAAGAAATCAAAAGTCATGTAACAACTTTGAATTAATTCATAT	5052
Db	1441	AATCCTTTAGTGAAGAAATCAAAAGTCATGTAACAACTTTGAATTAATTCATAT	1500
Qy	5053	ACAAAGCTTTCTGTGAGAGCGCTGCAAGTATTAAGCAGGCTGGAATCGAGAAATGTA	5112
Db	1501	ACAAAGCTTTCTGTGAGAGCGCTGCAAGTATTAAGCAGGCTGGAATCGAGAAATGTA	1560
Qy	5113	TCTGTCAATCGTGTCTGTGATGAACGGAAGCTTTAGTGAATTCGAGTTTGAAGA	5172
Db	1561	TCTGTCAATCGTGTCTGTGATGAACGGAAGCTTTAGTGAATTCGAGTTTGAAGA	1620
Qy	5173	GTAATCTTCTTCAATGTAATGCAAAAGTCAAAATTAATTAATTAATTAATTAATTAATTAAT	5232
Db	1621	GTAATCTTCTTCAATGTAATGCAAAAGTCAAAATTAATTAATTAATTAATTAATTAATTAAT	1680
Qy	5233	GCAAAATGAGAAAGCGGCTGAGTGGAGCAAGTGTCTATCAAAATTTTGAAGAA	5292
Db	1681	GCAAAATGAGAAAGCGGCTGAGTGGAGCAAGTGTCTATCAAAATTTTGAAGAA	1740
Qy	5293	CAATCAATTAATGCTATTGTAATAAATCAATAAATTAATTAATTAATTAATTAATTAATTAAT	5352
Db	1741	CAATCAATTAATGCTATTGTAATAAATCAATAAATTAATTAATTAATTAATTAATTAATTAAT	1800
Qy	5353	AAAAATTAATCAATGTAATGCAAAAGTATTAATTAATTAATTAATTAATTAATTAATTAAT	5412
Db	1801	AAAAATTAATCAATGTAATGCAAAAGTATTAATTAATTAATTAATTAATTAATTAATTAAT	1860
Qy	5413	CGAGAGCAAAAGAGCCCTGTGCAAGAGCTTCTGCAAGTACTTACCTTGAATAAGCA	5472
Db	1861	CGAGAGCAAAAGAGCCCTGTGCAAGAGCTTCTGCAAGTACTTACCTTGAATAAGCA	1920
Qy	5473	GTTTCTTCTCATGTTGATCAAACTGATATTGAACAAAGATTTAGAGAAAGAAATTAATGA	5532
Db	1921	GTTTCTTCTCATGTTGATCAAACTGATATTGAACAAAGATTTAGAGAAAGAAATTAATGA	1980
Qy	5533	AATAAGAAAGAGCAAAATGTTAATGTTCTAGCTGAAATTCAGAGTCAAGTGTCAAAAT	5592
Db	1981	AATAAGAAAGAGCAAAATGTTAATGTTCTAGCTGAAATTCAGAGTCAAGTGTCAAAAT	2040
Qy	5593	GCAGAGTGTCTTCCGAGAGCAAGTGAACAAGTGTGAGAGCTGTGAGTGAAGTAAAT	5652
Db	2041	GCAGAGTGTCTTCCGAGAGCAAGTGAACAAGTGTGAGAGCTGTGAGTGAAGTAAAT	2100
Qy	5653	AAAAATTAACAAATTAATCTTGTGACATATTAATAAATAATGATAC	5693

Db	2101	AAAAATTAACAAATTAATCTTGTGACATATTAATAAATAATGATAC	2141
RESULT 7			
AR442024			
LOCUS	AR442024	1887 bp	DNA linear PAT 20-FEB-2004
DEFINITION	Sequence 12 from patent US 6669940.		
ACCESSION	AR442024		
VERSION	AR442024.1	GI:4269175	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1887)		
TITLE	Negara, J. T. G., Stewart, G. C., Narayanan, S. K. and Chengappa, M. M.		
JOURNAL	Recombinant furobacterium necrophorum leukotoxin vaccine and		
PREPARATION	preparation thereof		
PATENT	Patent: US 6669940-A 12 30-DEC-2003;		
FEATURES	Kansas University Research Foundation; Manhattan, KS		
source	1. 1887		
ORIGIN	Location/Qualifiers		
Query Match	19.4%; Score 1887; DB 6; Length 1887;		
Best Local Similarity	100.0%; Pred. No. 7.4e-239;		
Matches 1887; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	5624	CTGCAGTAGAGAGCTGGAGTACAGTTAATAAATTAACAAATTAATCTTGTGACATATTA	5683
Db	1	CTGCAGTAGAGAGCTGGAGTACAGTTAATAAATTAACAAATTAATCTTGTGACATATTA	60
Qy	5684	AAATAGTACTCAAAATGTAGAAATGCTTTGGTATTAAGCAAAATCTTATCATCTATTA	5743
Db	61	AAATAGTACTCAAAATGTAGAAATGCTTTGGTATTAAGCAAAATCTTATCATCTATTA	120
Qy	5744	AAACAATGGAAATGGAGCTGAGTGGAGTGGAGCTGGAGAGCTGGAGTGAACAGTTCTGTA	5803
Db	121	AAACAATGGAAATGGAGCTGAGTGGAGTGGAGCTGGAGAGCTGGAGTGAACAGTTCTGTA	180
Qy	5804	CAGTGAATTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	5863
Db	181	CAGTGAATTAATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	240
Qy	5864	AGGAAATGTGGAGTTTACAGAGTGTGATGCGGTATTTGCTAATTAATTAATTAATTAAT	5923
Db	241	AGGAAATGTGGAGTTTACAGAGTGTGATGCGGTATTTGCTAATTAATTAATTAATTAAT	300
Qy	5924	TGTCTGAGTGGCCCGTGCAGCAATAGGAGCTCAACAGGTGTAATGAATTAACAGAT	5983
Db	301	TGTCTGAGTGGCCCGTGCAGCAATAGGAGCTCAACAGGTGTAATGAATTAACAGAT	360
Qy	5984	CTACAAAAGCATATGTAATAAATTTCTACAGTATTTCTTAAGAGAAACAGATGATTATA	6043
Db	361	CTACAAAAGCATATGTAATAAATTTCTACAGTATTTCTTAAGAGAAACAGATGATTATA	420
Qy	6044	TTACTACTCAAGGCAAGTGAATTAAGTGTGATTAAGTATTCAAAAATCTTAATATTA	6103
Db	421	TTACTACTCAAGGCAAGTGAATTAAGTGTGATTAAGTATTCAAAAATCTTAATATTA	480
Qy	6104	ACGAAGATTTCACAAAAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	6163
Db	481	ACGAAGATTTCACAAAAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	540
Qy	6164	CAGTACTCATATCTTTAAATCTTTAATGGCAAAATGCCGTGTTTCAAGACAGCCGAG	6223
Db	541	CAGTACTCATATCTTTAAATCTTTAATGGCAAAATGCCGTGTTTCAAGACAGCCGAG	600
Qy	6224	TGGCAGAACTGTTAATTAATCAACAAAGTTTATGAGAAACAGAAAGCTTGTGAGAAAT	6283
Db	601	TGGCAGAACTGTTAATTAATCAACAAAGTTTATGAGAAACAGAAAGCTTGTGAGAAAT	660

QY 6284 CTATATTAATGCAAAACATTATCTGTAAATCAGAGATTACAGCAATTCATCGAG 6343  
DB 661 CTATATTAATGCAAAACATTATCTGTAAATCAGAGATTACAGCAATTCATCGAG 720  
QY 6344 TAGTAGGTTCTGTGGTGTGTGTGAAATGAGAGTAGAGCTTCTTCTGATACATA 6403  
DB 721 TAGTAGGTTCTGTGGTGTGTGTGAAATGAGAGTAGAGCTTCTTCTGATACATA 780  
QY 6404 TTATATTAATGCAAAACATTATCTGTAAATCAGAGATTACAGCAATTCATCGAG 6463  
DB 781 TTATATTAATGCAAAACATTATCTGTAAATCAGAGATTACAGCAATTCATCGAG 840  
QY 6464 GAGAGAGCTGAATTTACAGAGATTCTTAAGCAAGAAATTCCTTTTGGAGTCGAG 6523  
DB 841 GAGAGAGCTGAATTTACAGAGATTCTTAAGCAAGAAATTCCTTTTGGAGTCGAG 900  
QY 6524 TCGCAGCAGCCGGGGTAGAGAGCCGAGAGTGCAGAGAACCGTTCCGTAATCAATTTGCAG 6583  
DB 901 TCGCAGCAGCCGGGGTAGAGAGCCGAGAGTGCAGAGAACCGTTCCGTAATCAATTTGCAG 960  
QY 6584 GAAAGACGAGATGATGTGGAAGAACAAAGATTTTGGTAAAAAAGCTGATTCAG 6643  
DB 961 GAAAGACGAGATGATGTGGAAGAACAAAGATTTTGGTAAAAAAGCTGATTCAG 1020  
QY 6644 CAAAACGTTATGTTCTGTGCAATTTGCAATTCGAGAGTGCAGAGAGAGAG 6703  
DB 1021 CAAAACGTTATGTTCTGTGCAATTTGCAATTCGAGAGTGCAGAGAGAGAG 1080  
QY 6704 CTGGAATTTGAGCAGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAG 6763  
DB 1081 CTGGAATTTGAGCAGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAG 1140  
QY 6764 AAAATTTCTAAATTTATGACTCGAAACAGTATGATTAATGAGAGAAATGAGATTAAT 6823  
DB 1141 AAAATTTCTAAATTTATGACTCGAAACAGTATGATTAATGAGAGAAATGAGATTAAT 1200  
QY 6824 CAGGATCTGGAATTCGTTTACGCGAGCTGGAATTTCTTACGCGAGATTCGAGAGTGC 6883  
DB 1201 CAGGATCTGGAATTCGTTTACGCGAGCTGGAATTTCTTACGCGAGATTCGAGAGTGC 1260  
QY 6884 TTTCTGTCATATATTTGCAAAATTAAGGTAGAAACAGATTCGATCTAGTACTTAACT 6943  
DB 1261 TTTCTGTCATATATTTGCAAAATTAAGGTAGAAACAGATTCGATCTAGTACTTAACT 1320  
QY 6944 CTTCTACTGATTAATGTTAAAGCTCTTAATTAATTTTGAATTCCTTACAGCGCGTG 7003  
DB 1321 CTTCTACTGATTAATGTTAAAGCTCTTAATTAATTTTGAATTCCTTACAGCGCGTG 1380  
QY 7004 GAGGAGCCGAGAGTCTTGCAGCAGTTCGCGAGTGTCTTCTTAACTATTAATTAAT 7063  
DB 1381 GAGGAGCCGAGAGTCTTGCAGCAGTTCGCGAGTGTCTTCTTAACTATTAATTAAT 1440  
QY 7064 CTGTGATGCTGAGATTCATTAATCTGATTTGAATTCCTTACAGAGAAATGTAATG 7123  
DB 1441 CTGTGATGCTGAGATTCATTAATCTGATTTGAATTCCTTACAGAGAAATGTAATG 1500  
QY 7124 TAAACGCAAAAGAGAAAAAATTAATTAAGCAAAACAGCAAAATGAGAGATTCGAGAG 7183  
DB 1501 TAAACGCAAAAGAGAAAAAATTAATTAAGCAAAACAGCAAAATGAGAGATTCGAGAG 1560  
QY 7184 CAGCAATCGAGAGCAATGCTTGTGTAAATTAATTTTGAACAGCTGTAAGAGATTAAGAAA 7243  
DB 1561 CAGCAATCGAGAGCAATGCTTGTGTAAATTAATTTTGAACAGCTGTAAGAGATTAAGAAA 1620  
QY 7244 ATTCTGAAGAAAGAGAAAGAGATTTTAAATCTTAAAGCAAGTAACTAAAGAGCAAG 7303  
DB 1621 ATTCTGAAGAAAGAGAAAGAGATTTTAAATCTTAAAGCAAGTAACTAAAGAGCAAG 1680  
QY 7304 ATTAAGAAAGTAAATGATGCTAGCAAAATTTTAACTAGCAGAGTATTTCTAGAGAG 7363  
DB 1681 ATTAAGAAAGTAAATGATGCTAGCAAAATTTTAACTAGCAGAGTATTTCTAGAGAG 1740

QY 7364 ATACTCTGTAAACCGAGTAGAGAGATCTAGAGGAGAAAGAAATTAAGCCATTGTGA 7423  
DB 1741 ATACTCTGTAAACCGAGTAGAGAGATCTAGAGGAGAAAGAAATTAAGCCATTGTGA 1800  
QY 7424 AGACTCTGTAAATTTTGAAGAAATTAAGTATTAATTAAGAGAGCAAGAAATTAATCA 7483  
DB 1801 AGACTCTGTAAATTTTGAAGAAATTAAGTATTAATTAAGAGAGCAAGAAATTAATCA 1860  
QY 7484 CTTCTACTGAGTGTGTGGAACTGCAG 7510  
DB 1861 CTTCTACTGAGTGTGTGGAACTGCAG 1887  
RESULT 8  
AR442021 1130 bp DNA linear PAT 20-FEB-2004  
LOCUS AR442021  
DEFINITION Sequence 9 from patent US 6669940.  
ACCESSION AR442021  
VERSION AR442021.1 GI:4269172  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
REFERENCE  
1 (bases 1 to 1130)  
Negaraja, T.G., Stewart, G.C., Narayanan, S.K. and Chengappa, M.M.  
TITLE Recombinant fuso bacterium necrophorum leukotoxin vaccine and  
preparation thereof  
JOURNAL Patent: US 6669940-A 9 30-DEC-2003;  
Kansas University Research Foundation; Manhattan, KS  
FEATURES  
source  
1..1130  
location/Qualifiers  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 11.6%; Score 1130; DB 6; Length 1130;  
Best Local Similarity 100.0%; Pred. No. 3.6e-119;  
Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGACGGGCTCAAAATTAACGTTCAAGAGCAAGAAAGAGATTCAGATTTCTAAAAA 60  
DB 1 ATGACGGGCTCAAAATTAACGTTCAAGAGCAAGAAAGAGATTCAGATTTCTAAAAA 60  
QY 61 GTTTAATGATTTTGGAGTTGTTGATTAACCTATGACGTGAGGCTTAATGATCAATC 120  
DB 61 GTTTAATGATTTTGGAGTTGTTGATTAACCTATGACGTGAGGCTTAATGATCAATC 120  
QY 121 ACCGCGACTGAATTTTGGAAACAAATTAAGAAAAAAGATTAATGATTAATGATTAAT 180  
DB 121 ACCGCGACTGAATTTTGGAAACAAATTAAGAAAAAAGATTAATGATTAATGATTAAT 180  
QY 181 ACAGCAAGATTCAGAGGAGAGACGCTTTTAAACGTTTAAATGATTAATGATTAATGAT 240  
DB 181 ACAGCAAGATTCAGAGGAGAGACGCTTTTAAACGTTTAAATGATTAATGATTAATGAT 240  
QY 181 ACAGCAAGATTCAGAGGAGAGACGCTTTTAAACGTTTAAATGATTAATGATTAATGAT 240  
DB 181 ACAGCAAGATTCAGAGGAGAGACGCTTTTAAACGTTTAAATGATTAATGATTAATGAT 240  
QY 241 AATAATTAAGCAAACTATATTTTGGGAGAAAGATTAAGTACGGGGTAAATATCTTTT 300  
DB 241 AATAATTAAGCAAACTATATTTTGGGAGAAAGATTAAGTACGGGGTAAATATCTTTT 300  
QY 241 AATAATTAAGCAAACTATATTTTGGGAGAAAGATTAAGTACGGGGTAAATATCTTTT 300  
DB 241 AATAATTAAGCAAACTATATTTTGGGAGAAAGATTAAGTACGGGGTAAATATCTTTT 300  
QY 301 AACTTTGTCATGAGAAAAATTAAGTAGATGATTAATCAAGCAATTCGAGAAAAATTA 360  
DB 301 AACTTTGTCATGAGAAAAATTAAGTAGATGATTAATCAAGCAATTCGAGAAAAATTA 360  
QY 361 ATTGAGAGAAATTAATTTCTTAAAGCTCGAAGGAGATGCGATGAGAAAAATGAGATT 420  
DB 361 ATTGAGAGAAATTAATTTCTTAAAGCTCGAAGGAGATGCGATGAGAAAAATGAGATT 420  
QY 421 ATCAATGCTGCTTTCTTCAATCTATTAATTCAGAAACAGATTAATTAAGAGGCTTG 480  
DB 421 ATCAATGCTGCTTTCTTCAATCTATTAATTCAGAAACAGATTAATTAAGAGGCTTG 480  
QY 481 GAAGAGCCAAACATGTTAAAGTTTAAATGATCAATTCAGATGAGAGAGAGAGAG 540  
DB 481 GAAGAGCCAAACATGTTAAAGTTTAAATGATCAATTCAGATGAGAGAGAGAGAGAG 540





Db 4596 TAGCAGTACGAGTACGAGAACTGTAGCAGTACGAGAACTGTAGCAGTACGAGCAG 45937  
 Qy 7082 ACAATTAATCTGATTTGACTTCGTAACGAGAAAAAGTAATGTAAACGAGAAAAAGAGAAA 7141  
 Db 45936 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45877  
 Qy 7142 AAAATATTGAACAAACAGCAGCAA--TGCAGGAATCGAGAGAGCAGCAATCGAGCCA 7198  
 Db 45876 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45817  
 Qy 7199 ATGCTTGGTAAATATTTTGGAAACGCTGTGAAGATAGAAAAAATTCGTAAGGAAAAAG 7258  
 Db 45816 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45757  
 Qy 7259 GAACAGAACTTTTAAAACTTTTGAAGAAATTAAAGAAACAGATTAATAAAAGTAAATG 7318  
 Db 45756 TAGCAGAACTGTAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45697  
 Qy 7319 ATGCTACGAAAAAATCTTCAATCAGCAGTATTTCTACAGAAATCTTCTGTAAAG 7378  
 Db 45696 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45637  
 Qy 7379 CGGATAGAGAGTACTCAGAGGAGAAAGAAATTAAAGCCATTGTGAAGACTTCTGTATTA 7438  
 Db 45636 TAGCAGAACTGTAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45577  
 Qy 7439 TTGAAAAAATGTAGATATTACCAACAGAGCAAGAAATATATCATCTTCTAGTGTGTT 7498  
 Db 45576 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45517  
 Qy 7499 TGGAACTGCAAGTCTTCTCCGCTCCGCTACGAGAACTGCAAGTTCACAAATTTAAAAAGAA 7558  
 Db 45516 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45457  
 Qy 7559 ATTCGGAATTAATCTGTTGAAAAATCTTTGTGAAAGACCTGAAAAAGTAAATGTAGAT 7618  
 Db 45456 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45397  
 Qy 7619 CGGATTTACAGAAATGTTGCTTTTAAACAGCATATCAAGGCTCTGTAGAGCATTGGGAA 7678  
 Db 45396 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45337  
 Qy 7679 TAGGAGCTGCTTACGAAATTAATTTCTAATGGAAGATCAAAATATCAATTTAAAAATT 7738  
 Db 45336 TAGCAGAACTGTAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45277  
 Qy 7739 CTAAGCTATTAGGAAAAAATTTGATGTTATTGTATAAATAATTCGGAAATTTGAAGCCG 7798  
 Db 45276 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45217  
 Qy 7799 AAGCAAAAAGATTTAACCTGAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAAGA 7858  
 Db 45216 AAGCTGTACAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45157  
 Qy 7859 ATGAATTAATTCAGAGGTTGAAATTGAGAAGATATTTTCAATGAAGAAAAATAGATTA 7918  
 Db 45156 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45097  
 Qy 7919 CTAAGCCCTCTAAAGAAATTTGAAGAAATCAATCTCAAGTGAAGAAAAAGAAAAACAGG 7978  
 Db 45096 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 45037  
 Qy 7979 TGAAGTCTGAATCTCAAGAGCTTCTGTAGAGCAGTACGAGGAGGAGAAATTTATTTCCG 8038  
 Db 45036 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 44977  
 Qy 8039 AAGCAAAAGATGCCGAGAGCTTTATTGAAAGTTAGTACAAAATCCGAGAGAAATATT 8098  
 Db 44976 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 44917  
 Qy 8099 TTCATGACATATAATGTGAATATGGAAGCAACATATAAATGAAAGTAAACAGCAGTTTCTA 8158  
 Db 44916 AAGCTGTACAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 44857

Qy 8159 AAGCAGTACAGGTTCTGTATTGGAGAGATTGAGTACCAAGCAGAGAGTACTGCTG 8218  
 Db 44856 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 44797  
 Qy 8219 CAGTAAAACTGTGTAGTAAAGTTGAGAAAGAAATTTGTTCAAGAACTGATGTAATG 8278  
 Db 44796 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 44737  
 Qy 8279 CAATTTCTAAGTACAGGTTTGGATGAAAGATTAAGTAACTGCTAAATCTTGTATGAT 8338  
 Db 44736 TAGCAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAG 44677  
 Qy 8339 CAGCAAAATGAGAGAAATTGCCGAGAGAGAGTGAATTAATCTTCAAGCA 8388  
 Db 44676 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 44627

RESULT 10  
 AC024253/C  
 LOCUS  
 DEFINITION  
 SEQUENCE, 12 unordered pieces.  
 ACCESSION  
 AC024253  
 VERSION  
 HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 179553)  
 Wacreston,R.H.  
 The sequence of Homo sapiens clone  
 2 (bases 1 to 179553)  
 Wacreston,R.H.  
 Direct Submission  
 Submitted (28-FEB-2000) Genome Sequencing Center, Washington  
 University School of Medicine, 444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 On Sep 1, 2000 this sequence version replaced gi:8568958.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

## COMMENT

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: MUGSC  
 Web site: http://genome.wustl.edu/gencl/index.shtml  
 Project information -----  
 Center project name: H.NH0758C19  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 100%  
 Sequencing vector: plasmid; 0%  
 Chemistry: Dye-terminator Big Dye; 0% of reads  
 Chemistry: Dye-terminator Big Dye; 0% of reads  
 Assembly program: Phrap; version 0.950319  
 Consensus quality: 175235 bases at least Q40  
 Consensus quality: 176522 bases at least Q30  
 Consensus quality: 177171 bases at least Q20  
 Insert size: 19800; agarose-fp  
 Insert size: 178453; sum-of-contigs  
 Quality coverage: 4.48 in Q20 bases; sum-of-contigs  
 Quality coverage: 4.99 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 7554: contig of 7554 bp in length  
 \* 7555 7654: gap of unknown length



Qy 6175 ACTTTAAATCTTTATTGCGAAATGCGGTGTCAGACAAAGCCGAGTGCAGAACT 6234  
Db 175564 GCACACGACACAGCAGAGTAGCAACAGAAATCGAGTAATTAACAGTAGCAGCAGACA 175505  
Qy 6235 GTTAATATCAACAGGTTTATGAGAAACAGAAAGCTCTTGTAAGAAATTTCTATTAAAT 6294  
Db 175504 GCAGCAATCATTTCCATAGTAGCAACAGATAGCAGCAACCCAGCAGCAATATGAGA 175445  
Qy 6235 GCAAAACATTTATCTGTAAATATCAGAGATTACAGCAATTCAGAGTAGAGTTCT 6354  
Db 175444 GCATATGACACAGATATAGTAGCAACAGCAATAGCAACAGTAGCAGCATACACACA 175385  
Qy 6355 GTTGCTGTTGTGGAATGTAGAGTAGAGCTTCTTGATACCAATATTATTAAGA 6414  
Db 175384 ACAGCAACAGCAGCAGCAACATTAAGACAGCAGCAACAAACAGTAGCACAGCA 175325  
Qy 6415 AATACCAAGACAAAGTTGGAATACTACATGTCGATGAAGGTTTCGGAATAAGT 6474  
Db 175324 GCACACACAGCAACATTAACAAAGCAACGATAGCAACGATACAGCAACAGCAACAGA 175265  
Qy 6475 GAATTTACAGCAGATTCTAAGCAAGAAATTTCTCTTTGAGTGCAGAGTCGACAGCC 6534  
Db 175264 TCAGCAACACAGCAGCAACATACAGCAGTAAACAGCAACAGCAGCAGCAACAGCA 175205  
Qy 6535 GGGGTAGAGCCGAGTGCAGGAAACGTTTCCGTAAATCAATTTGCAAGAAACGGA 6594  
Db 175204 ACAGCAGCAGCAACATTAAGACAGCAACAGCAACAGTAGCAGCAGCAACCAACA 175145  
Qy 6595 GTTAGATGTGGAAGAAAGCAAGTTTGTGTAAGAAAGCTGAAATTAACAGCAACAGTTAT 6654  
Db 175144 GCACGCAACAAAGCAACAAATAGCAACGATACAGCAACAGCAACAGCAATACCAACA 175085  
Qy 6655 AGTTCTGTTGCAATTTGGAATGCGGAGTGCAGAGTGCAGCAAGAGC---TGGAAAT 6711  
Db 175084 AAGAGATCAGCAACAGCATATAGCATCAATTAACACAGCAACAGCAGCATACACA 175025  
Qy 6712 GAGCAGCAGTGCAGCTTACCAAGATGATCAACACAGCAAGCAAGTGAATAATCT 6771  
Db 175024 GAATTCACAGTAACACAGTAGCAGCAGCAATTAATTCATAGTAGCAACCAACA 174965  
Qy 6772 AAAATATATGCTCGAAACAAAGTTAGTGTATAGCAAGAAATAGATTAATACAGTACT 6831  
Db 174964 GTAGCAGCAACACAGCAGCAGTAGCAGCAACAGTAATAGTAGCAGCAACA 174905  
Qy 6832 GGAATCGTTGACCCGAGCTGGAATTTCTTGACCCGAGATATCTGGAATGTTCTGTC 6891  
Db 174904 GCATATGCAACAGTAGCAGCATCAGCAACAAAGCAACATAGCAGCATTAACA 174845  
Qy 6892 AATATATTTGCAATTAAGTTAGAAACAGATATGATCATAGTACTTACTTCTACT 6951  
Db 174844 GCACAGCAGCAGCAACATTAAGACAGCAGCAACAAACAGTAGCAGCAGCA 174785  
Qy 6952 GATGTAAATGTAAAGCTCTTAATTAATTTTGAAATTCCTTGACCCGCTGGAAGAGC 7011  
Db 174784 ACAACGCAACATTAACAAAGCAACGATAGCAACAGTATCAGCAACAGCAACAGCAAT 174725  
Qy 7012 GCAAGCTTTCGACAGTACCGAGTGTCTTGTTACATTAATTAATAGTCTGTGATA 7071  
Db 174724 GCACAGCAGCAGCATCAGCAACATAGCAGATTAACAGCAACAGCAGCAGCAACA 174665  
Qy 7072 GCTCGAGTTCACAATTACTGTGATTTGACTTCGTGAGCAAGAAAGTAATGTAAACGCA 7131  
Db 174664 ACAGCAACAGCAGCAGCAACATTAAGACAGCAGCAACAGCAACAGTAGCAGCAGCA 174605  
Qy 7132 AAAAGGAAAAAATTTTAAGCAACAGCAGCAATGCGAATGCGAGAGCAGCAATC 7191  
Db 174604 ACACAGCAACATTAACAAAGCAACGATAGCAACAGTATCAGCAACAGCAACAGCAAT 174545  
Qy 7192 GAGGCAATGTCTTGTTAATATTTTGAACAGCTGTAGAAATGAAAAATTTCTGA 7251  
Db 174544 GCACACAAAGCATCAGCAACAGCAATAGCATCAATTAACACAGCAACATAGCACA 174485  
Qy 7252 GGAAGAGAACAGAGTTTAAACTTTAGACGAAGTTTACAAAGAACAGATTAATAA 7311

Db 174484 GCACAGCAACAGCAGCAGTAGCAACAGAAATTCAGTAATTAACAGTAGCAGCAGCA 174425  
Qy 7312 GTTAATGATGCTGAGAAAAAATCTTACATCAGAGGATTTCTACAGAAATACTCT 7371  
Db 174424 GCATCATTTACATATGTAAGCAACAGTAGCAGCAACAGCAGCATATGAGCAGCA 174365  
Qy 7372 GTAAAGCGATTAAGAGATCTCAGGAGAAAGAAATTAAGCCATTTGTGAAGCTTCT 7431  
Db 174364 ATAGCAACAGTAATGTAGCAGCAACAGCAATAGCAACAGTAGCAGCATCAGCAACACA 174305  
Qy 7432 GATATTATGAAAAATGTAGATTTTCAACAGAGCAAGAAATATATCATCTTACT 7491  
Db 174304 GCACAGCAGCAGCAAC-----ATAAGACAGCCGAGCAACAGCAACAGTAGCAGCA 174251  
Qy 7492 GTTGTTGGGACGTGACAGTCTTGCTTCGCGATCAGAGCAGTAGGAGTTACAAATTT 7551  
Db 174250 GCAGCAGCAACAGCAGTAGCAGCAGCAACAGCAACAGCAACATTAACAAAGCAACAT 174191  
Qy 7552 AAAAGAAATTCGAGATTACTGTGAAATTTCTTTGAAAGCAGCTGAAAAAGTAAT 7611  
Db 174190 GCACAGATCAGCAACAGCAATAGCAGCAGCATCAGCAACAGCAACATATAGCA 174131  
Qy 7612 GTTAGATGATTTATACAGAAATGTTGCTTTAAGCATATCAAGTCTGTAGAGCA 7671  
Db 174130 GTAAACACAGCAACAGCAGCAGCAACAGCAACAGCAGCAGCAACATTAAGAGCA 174071  
Qy 7672 TTGGGAATAGAGCTGCTTACAGAAATTAATTTGAATGGAAGTCAATATCAGTATT 7731  
Db 174070 ACACAGCAGCAACAGTAGCAGCAGCAGCAACAGCAACATTAACAAAGCAACATAGCA 174011  
Qy 7732 AAAATTTAAGCTATTAGAAAAAATTTATGATTTATTTGAAGATTAATTCGAAATTG 7791  
Db 174010 GTATCAGCAACAGCAACAGCATATGAACAAAGATCAGCAACAGCAATAGCATCA 173951  
Qy 7792 AGAGCGAAGCAAAAGATTAAACGTAGAGCGTGTGCTGCGAGCAATATCTCAAA 7851  
Db 173950 TTAACACAGCAACATGTGCCAGCAGCAACAGCAACAGCAGTAGCAACAGAAATCA 173891  
Qy 7852 GCAAGAAATGAATGAATTTAGAGGTTGAATTTGAAGAGATTTTCAATGAAGAAAT 7911  
Db 173890 GTATTAACAGTAGCAGCAGCAGCAGCAGCAGCAACATTTCCATAGTAGCAACACA 173831  
Qy 7912 AGAGTACTAGCCCTTTCTTAAGAAATTTGGAAGAAATCAATGTCAAGTGGAAAAAGAA 7971  
Db 173830 GTAGCAGCAACCCAGCAGCATATAGTAGCAATAGCAACAGTAATAGTAGCAACACA 173771  
Qy 7972 AACAGATGACTGCTGAATCTCAG-----GAGCTTCTGTAAGAGCAGTAGCAGGCA 8025  
Db 173770 GCATATGCAACAGTAGCAGCATTCGCAACAAAGCAACATAGCAGTAGTAACACAGCA 173711  
Qy 8026 GGAATTTATTCGAAGCAAAAGATGCGGAAGCTTTATTTGAAGTTAGTAACAATCC 8085  
Db 173710 ACAGCAGCAGCAACATATGAGAGCCCGCAGCAACAGCAACAGTAGCAGCAGCAACA 173651  
Qy 8086 GGAAGAAATTTTTCATGACAGATATATGATATGAAAGCAACATTAATGAAGTA 8145  
Db 173650 ACAGCAACATTAACAAAGCAACAGTAGCAACAGTAATCAGCAACAGCAATAGCA 173591  
Qy 8146 ACAGCAGTTTCTAAAGCAGTAACAGTCTGTATTGGAGAGATTGGAGTCAACAGCA 8205  
Db 173590 ACAGCAGCATCAGCAACAGCAATAGCATCAACATTAACAGCAACATTTAGCCACACA 173531  
Qy 8206 GAAGCTACTGCTGCAAGTAAC-----TATGTAAGAGTTGAGGAAGAAAT 8253  
Db 173530 ACAGCAACAGCAGCAGTAGCAACAGAAATAGCAGTAATAGCAGCAGCAGTAGCACA 173471  
Qy 8254 TTGTTCAAGCAACATTCAGATTAGATCAATTTCTAAAGTAGAAGGTTGAGTAGAATTA 8313  
Db 173470 ATTAATTAACATAGCAGCAATAGCAGCAACAGCAACAGTAGTAAGAGCAACAAATTA 173411  
Qy 8314 GTACTGCTAAATCTTCTGTAGTATCAGAAATGAGAGAAATGCGAGCAGAGTG 8373

Db		173410	AAGCAACAACGACGCAATACGACGAATCATTAACGCAATAGCAGCAGAACA	173351
Oy		8374	AATACTTCCAGCACAAAGTAATACTGAATCGGTATTGGTTTAGCAAAGCAAGTTAT	8433
Db		173350	ACGCAGATPAACAATAGCAGTAATGACAAATCATCAGCAACGATGACAAACAGTGTCAT	173291
Oy		8434	GAAAATTAATGATTACCAAAAAAATATATTTCAGAAAGTCGAATGCTTGCTTTAAATGAT	8493
Db		173290	GCAACAGTAGCAGCAGTGCACCAATATGCAACAGCATCATTAATCAACATGGCAATPACA	173231
Oy		8494	ACAAAGAAATGAAGCGAAATTTAGAAATCTTTTAGCGCGTACCGBTGTCATCACAAAGAAC	8553
Db		173230	ACGAAAGCAGCAGCAACAAACAAAGCAACATAATGACGAATATGACAGAACACACAAC	173171
Oy		8554	AACAAGCATTTATCGAGATCAAAACAAGTTAATCTTCTTACCAACTGTATAATGAGGAAAACT	8613
Db		173170	GTPACAGTAGCAGCAACATCAGCAGATGATPACACACAGCAGTTGTCAGTAAATATGTTG	173111
Oy		8614	TCTCAACTTCGTCGCAAAAGCTTTGGCTAAATAATGAAATTAATGGAATATGAAAGAACT	8673
Db		173110	ACAACAGCAGCAGCAGCGCCATAGTAGTAGCAATGACAAACAGCATTAAACMACAGCAACA	173051
Oy		8674	GGAAGGCCCTTAGTGCAGCGGAAACAGCAGCCGTTGAAATAATTAACAA	8722
Db		173050	ATTTCAGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	173002
RESULT 11				
PFAMESA				
LOCUS	PFAMESA	4995 bp	mRNA	linear INV 26-APR-1993
DEFINITION	Plasmodium falciparum (MESA) mRNA exons 1-2, complete cds.			
ACCESSION	M69183			
VERSION	M69183.1 GI:160408			
KEYWORDS	erythrocyte surface antigen.			
SOURCE	Plasmodium falciparum (malaria parasite P. falciparum)			
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
REFERENCE	1 (bases 1 to 4995) Coppel, R.L. Repeat structures in a plasmodium falciparum protein (MESA) that binds human erythrocyte protein 4.1 Mol. Biochem. Parasitol. 50 (2), 335-347 (1992)			
AUTHORS	JOURNAL	PUBMED	COMMENT	
			Original source text: Plasmodium falciparum (strain D10) cDNA to mRNA.	
FEATURES	source	Location/Qualifiers		
		1..4995	/organism="Plasmodium falciparum"	
			/mol_type="mRNA"	
			/strain="D10"	
			/db_xref="taxon:5833"	
			jctn(278..379..502.. .4932)	
gene			/gene="MESA"	
CDS			jctn(278..379..502.. .4932)	
			/gene="MESA"	
			/codon_start=1	
			/product="mature-parasite-infected erythrocyte surface antigen"	
			/protein_id="AA29651.1"	
			/db_xref="GI:160409"	
			/translation="MENEGNKVKVYNNSSLKYYMKFCLTIICVFLLDIYNCSPT VYSILIKNNDRIVRLISETPEPMSELEIRTEPEDDLVSIIRNYIECLRNPYIDPL WGSVTDKNDKROHIT KLDEMSEERRKKOEBNADIDEIRKEKGYLIMKEEMD ESNVEKAPELQRIKLARDPRPKHYVMESKETGESKTDESKTESKETGESKE TGSGKETGESKETGESKETGESKETGESKETGESKETGESKETGESKETGES KETGESKETGESKETGESKETGESKETGESKETGESKETGESKETGESKETGES ESQRDBGNKVSGPYENSNSNVTSSESLETYKLAKEBENBEKLEVNNDGASNSBDP KKTLQBENGSTKESSEETKDQDPRENEKKAADNNKKSFKKKSPFQM/GCNFLCKNTL TGDDEBTLYVDKDAKKKHGFALREANTKNENEKKDKLVIGEDGDEVKEKDEOKVL DEBEDKDEVKENDBKDKVLGEEDKENVEKNGEKDKXVI GSEKTOEIKERKYRYK KKKKYYKKCIKENDTBGDNKVKAPEIIITEVEBEIKOVJEDGINKENDTBGDNVAKP EIIITEVEBEIKOVJEDGINKENDTBGDNKVP IITEVEBEIKOVJEDGINKENDTB GNDVKVPEIITEVEBEIKOVJEDGINKENDTBKDVIOEIITEVEBEIKEQEERK	

GNNKTEILIKDVIIGQEVIIIEYKVIKKVVEGIIKENHDSOKYIGQIIIEYVEV  
 EIKQIYBEGIKENDTSSKDVIGIIEVIMVEEGENDOKYTKOKKVEYKVEKK  
 VKKRVKRNKRNKRNKDVIGKEIMKEDVNEKDTANDKIEQKEKEVEYKVEYK  
 EYKVEKEVEYKVEKEVEKEVEKEVKEKOTSKDIEIQEKEVEYKVEKOTEN  
 KDVIVIGQIIIEIKKVKRVKRVKRNKRNKONVIVQIIMNEDVNEKDTANDKYE  
 QKEKEVEKEVEKEVEKEVEKEVEYKVEKEVEKEDNESDNTVQIIMNED  
 VNEKTESKDMIGKEVYIIEYKVEYKRVKRNKRNKRNKRNKDVIEQIIVSE  
 EVNEKOTENKDKIKRVKRPIDDKKREVOIIMNEDVNEKTESKDMIGKEVYIIE  
 EYKVEYKRVKRNKRVKRNKRVKRNKRNKRVYIIEQIIVSEVNEKOTENKDKIKRVK  
 PIDDKKREVOIIESESESESESESESESESESESESESESESESESESESE  
 SE  
 QNEEYIPHEEYIIOQKHNPLNTPNNAIANDPREPFINIISHEDTLMELKILVNL  
 FLEYNPHNNKQKAEIDKKFSENNVHQIILNTEYERIRMYEENARGLNLTVIDPIIT  
 FNVIYFGDDTWFKFIDE"

exon  
 /gene= "MESA"  
 /number=1  
 380. . 501

intron  
 /gene= "MESA"  
 /number=1  
 502. . 4932

exon  
 /gene= "MESA"  
 /number=2

ORIGIN

	Query Match	1.3%:	Score 123:	DB 2:	Length 4995:
	Best Local Similarity	43.6%:	Pred. No. 6.3e-07:		
	Matches 1032:	Conservative	0:	Mismatches 1230:	Indels 45: Gaps 9:
Qy	7110	AGAAAAAGTAATGTAA	CGCGCAAAAAGGAAAAAAATATTTAAGCAAA	CAGCAGCAATTC	7169
Db	1482	AGAGACCAAAAATTTAC	CCGAAAAAGAGAAATGAGGAGAAAAAATTAGACAAAAATGT		1541
Qy	7170	AGGAATCGGAGAGCAGCA	ATCGGACCAATGCTTGTAATTAATTTGGAA	CAGCGT	7229
Db	1542	TAAATGATGGGCAATCA	AAAAATTCAGAAATCC-----AAAAATTAACAGAACAGA		1595
Qy	7230	AGAGATAGAAAAAATTC	TGAAGAAAAAGAACAGAAATTTAAAAA	CTTTAGA	7289
Db	1596	AGAAAAATGTAACAAAG	AAAGTTCTGAAGAAACAAATGATTAACCGAAGAAAAATGA		1655
Qy	7290	TAAACAAGACAAATTA	AAAAAGTAAATGATGTAACGAAAAAAATCTTACA	ATCAGCAG	7349
Db	1656	GAAAAAGCGAGTATATA	AAAAAAGCTAAAAAAGAAAAAATCATTTTCAAATGTT		1715
Qy	7350	TATTTCTACAGAAATAT	CTCTGTAAGGCGATAGAGAGATCAAGGAGAGAAAT		7409
Db	1716	AGGATGTAATTTCTTA	TGTATTAATAATTTGA---AACTGATGATGAAGAAACGTT		1772
Qy	7410	TAAAGCAATGTGAAG	CTTCTGATATTATTGAAAAAATGTAGATTTACA	CAAGAGA	7466
Db	1773	GGTGTAAAGATATGCTA	AAAAAGAAACATAAAATTTTAAAGAAAGCTAATACTGAAA		1832
Qy	7470	CAAGATATATCACTT	CTACTGATGGTTTGGGAAGTCCAGAGCTTGCTTC	CGCATCAG	7528
Db	1833	AAATGATATATAAAGAA	GAATATAATTTATTTGGAAGAGAGATTAAGAGATGTTAAAGA		1892
Qy	7530	AACAGTGCAGTTACA	ATATTTAAAGAAATTC	CGAGTTACTGTTGAAA	7589
Db	1893	AAAGATGATGAACGA	AAAGATTAAGATTTAGGAAGAGAGATTAAGAGATGTTAAAGA		1952
Qy	7590	GAAAGCAGCTGAAAA	AGTAAATGTTAGTCCGATTTACAGAAATGTTGCTTTAA	CAGC	7645
Db	1953	AAAGATGATGAACGA	AAAGATTAAGTTTAGAGAAAGCAATTAAGAAAGTGTAAAGA		2013
Qy	7650	ATATCAAGGCTCTG	AGAGCATTTGGGAATGAGAGCTGCTATCGAATTTAAATTC	TAA	7705
Db	2013	AAAGATGATGGAAGA	AAAGATTAAGTATGGAATCAGAAAAAACACAAAGGAATTTAA		2072
Qy	7710	TGGAAGATCAAAAT	TACGATTTAAAAAATCTTAAGCTATTAGAAAAAATATGATGTTAT		7767
Db	2073	AGAAAAAGTAAAAA	AAAGATTTAAAA---AAAGTAAAAAAGTAAAGTAAAAAAGAAAT		2125

QY	7770	AGTAAAGATTAATCGGAATTGAGCGGAAAGCAAAAGATTAAACCGTAGGCGGTGAC	7822
Db	2130	TAAAGAAATGATCTGAAAGCTAACGATTAAGTGAAGGACCGAATAATTAATTGGAAG	2188
QY	7830	TGCCGAGCCATTATCTCAAAAGCAAGAAATGAAATGCAATTCAGAGTTGAAATTGAGA	7888
Db	2190	AG-----TAAAGAAAGAAATTTAAAAACAAGTAGAAGATGGAATTTAAAG	223
QY	7890	GAGTATTTCAATGAAGAAATAGGTAACTAGCCCTTCTAAAGAAATTGGAAGAAAT	794
Db	2235	AAATGATCTGAAAGGTAAACGATTAAGTGAAGGCGCAGAAATTAATTAACGAAGGTAAA	229
QY	7950	CAATGCTAAAGTGGAAAAAGAAAAAGAGTGACTCTGTAATGTCAGAGAGCTTCTGTAG	800
Db	2295	AGAGAAATTTAAAAACAAGTAGAAGATGAAATTTAAAGAAAAATGTAAGTAAAGCA	235
QY	8010	AGCAGTAGCAGGCGCAGGAATTTATTTCCGAAGCAAAAGATGCGGAGACTCTTAATTTGA	806
Db	2355	TAAAGTGAAGGGCCAGAAATTAATTAACGTGAAGAACTTAAGAGAAATTTAAAAACAAGT	241
QY	8070	AGTTAGTACAAATCCGAGAGAGATTTTTCATGACAGATTAATGGAATATGAAAGCAAC	812
Db	2415	AGAGAAAGGAATTTAAAGAAATGATACTGAAAGGTAAACATTAAGTGAAGGCGCAGAAAT	247
QY	8130	ACATTAATGAAAGTAAAGCAGC---TTCTTAAGACATTAACAGTTCTGTATTGGAG	818
Db	2475	AATATCTGAGAGTAAAGTAAAGAAATTTAAAAACAAGTAGAAGAGAGATTTAAAGAAA	253
QY	8187	AGTTGAGATCAACAGCGAGAGCTACTGCTGACAGTAAAACTAATGTAGAGAGTTGAGGA	824
Db	2535	TGATATCTGAAAGTAAAGATTAAGTGAATAGACAAATAATTAATCTGAAGAGATTAAGAA	259
QY	8247	AGGAAATTTGTTCAGACAAATTCGATTGAATGCAATTTCTAAAGTAGAAGGTTGATGA	830
Db	2595	AGAAATGMAAAACAGAGAAAGAAAAAGAAATTAAGAAATATCTTTGAAATTTAAAGATAT	265
QY	8307	AGATTAAGTAATCTGTAATCTTCTGTAGTATCAGAAATGAGAGAGAAATTCGCCGAGC	836
Db	2655	AGTATTTGACAAAGAGTAAATTAATTAAGAAAGTAAAGAAAGTAAATTAAGAAAAAGTGA	271
QY	8367	AGAGTGAATCTTCTTAACAGCAACAAGTAACTACATCCGATGTCGTTTACGAAGCA	842
Db	2715	AAAAGGAATTTAAAGAAATTCATCTGAAGAGTAAAGATTAAGTGAATAGACAAAGAAATAT	277
QY	8427	AGATTATGAATAATATGATTAACAAAAAAATATATTTCAGAGTCAATGCTCTTCTT	848
Db	2775	AGTGAAGAAATTAAGAAAGAAATTTGAAAAACAAGTAGAAGAGAAATTAAGAAAAATGA	283
QY	8487	AAATGATACAAAGATGAAGGAAATATAGATCTTTACCGGTAGCCGTTGATGCA	854
Db	2835	TACTGAAGTAAAGTAAAGTATAGGAGAAAGTAAGTAAAGGAGATGTTAATGGA	289
QY	8547	AGGAACAACAAGCATTTACGAGATCAACAAAGTTAACTTCAACCTGTAAATGAG	860
Db	2895	AGGTCCGAAACCAAGATTAAGTGAACAAACAGAAAGTAAAGAAAGTTAAAGAAAG	295
QY	8607	AAACGTATCTCACTTCGTGCAAAAGCTTTGGCTAAATATGAAATTAATGAAATGTAAA	866
Db	2955	AGTAAAAAAGAAAGTTAAAAAAGAGTAAAAAAGAAATATTAAGATGAAA---GAA	301
QY	8667	AGGAAGTGGAGAGGCTTATGTCGAGCGGAAACAGACCGTTGAAAATTTACAAAG	872
Db	3012	AGATTAATGTGATAGGAAAAAGAAATTAATGAAGAAATGTTAATGAAAAAGTATCCGCA	307
QY	8727	TACTACAGAGCATGTTGTCAGGAAATTTGGAAATTTGAGATTAATTAAGAAAGATTC	878
Db	3072	CAAAGATTAAGAGATGAAACAGAAAAAGAAAAAGAGAGTTAAGAAAAAGAAAGT	313
QY	8787	AAGAGATTAATGCAATTTGTAGAGTCAACGAGACGGAACCAAGAGGCTTTGTTCGAAA	884
Db	3132	TAAAGAAAAAGAGAGTAAAGAAAAAGAAAGAGTAAAGAAAAAGAAAGTAAAG	319
QY	8847	GAAATGATTTTCTGTGAAAAATACATTTCAAGGGAAGCAAAATATCATTCATTGAAGTAA	890

Db	3192	AAAAAGAAAGTAAAAAGAAAAAGAAAGTAAAAAGAAAAAGTAAACCGAAAGCAAAAGTTAA	3251
Qy	8907	AGCCGAAATGTTGGAAACCGGAAGTGTAAATGTAGATGCTTGAATGAACCTTGAATGTAGA	8966
Db	3252	AGAGATGTGAACCAAGAAAAAGAAAAAGAAAGTAAAAAGAAAGTTAAAGAAAAAGTACCGA	3311
Qy	8967	TCCTCAAGGAAAAAGTGTGGCTGTGTGGAAAT-----TGCTATTGGAAATGTTGATGT	9020
Db	3312	AAACAAAGATAAAGTGTAGGACAAAGAAATATATAGAAAGAAATAAAAAGAAAGTTAA	3371
Qy	9021	AAATTAATGTATTAAGAAAAATGTAGAAGCCAAATATGGAAGACATGCTATGTGAGAAC	9080
Db	3372	AAAAAGAGTAAAAAGAAATTAATAAAAATGAAAAACAAAGATATGTGATGATCAAGA	3431
Qy	9081	TACTGAAAAACAAGAAATATCAGACATTTTACAAGAGCAAAAGTAAATATTCCTGGAAAAAG	9140
Db	3432	AATATATGACGAGATGTGTAACGAAAAAGATACCGCAACAAAGATTAAGGTGATGAAACA	3491
Qy	9141	AGACGCTGCAGCTGCAGCTGCATATGCAATGTACACATT---TCCATGAGATGATGA	9196
Db	3492	AGAAAAAGAAAAAGAAAGATTAAGAAAAAGAAAGAAAGTAAAGAAAAAGAAAGAACTTAA	3551
Qy	9197	TTAAAAATTTGGCAACACGATATGCATCTTCTCAATTAATACCAAAATTCAAAAATA	9256
Db	3552	AGAAAAAGAAAGATTAAGAAAAAGAAAGAAAGTAAAGAAAAAGAAAGTAAAGAAAA	3611
Qy	9257	ATATTACTTTAGCATCAAGTATGTAATCGAATGTGA--TGTTCAATGGGTGGCTGGAAC	9314
Db	3612	AGATATCCGAAAGCAGAGATATATGTGATATGATCAAGAAATATGAAACGAAGATGTTAACGA	3671
Qy	9315	AAGAGTGCAGAGACCAGACGACATTTAGTGTGAAGATTCAAATTAATAGAACTAATA	9374
Db	3672	AAAAGATATCCGAAAGCAAAAGATTAAGAAATGAAAGAAAAAGAAAGTAAATTAAGAAAGTAA	3731
Qy	9375	TGTTGATTTAGCAGGAAAAATTTAAACAGAGGGGAAACATCAATGTATATGCCGATATGA	9434
Db	3732	AGAAAGAGTTAAAAAGAGATTAACAAAGAGTTAAACAAAGAGTAAACAGAAAGATAG	3791
Qy	9435	TAAAAATTTATATATTAAGTAAAGCAAA	9461
Db	3792	AAAAAATGAAAGAAAGATGTGATGCA	3818
RESULT 12			
AE014843			
LOCUS	AE014843	271546 bp	DNA linear INV 07-OCT-2002
DEFINITION	Plasmodium falciparum 3D7 chromosome 11 section 8 of 8 of the complete sequence.		
ACCESSION	AE014843	AE014186	
VERSION	AE014843.1	GI:23496380	
KEYWORDS			
SOURCE	Plasmodium falciparum 3D7		
ORGANISM	Plasmodium falciparum 3D7		
REFERENCE	1 (bases 1 to 271546)		
AUTHORS	Gardner,M.J., Hall,N., Fung,B., White,O., Bertrman,M., Hyman,R.W., Carlson,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyte,S., Chan,M.-S., Nene,V., Shallow,S.J., Suh,B., Peterson,J., Angioli,S., Pertea,M., Allen,J., Selengut,J., Haft,D., Mather,M.W., Valiya,A.B., Martin,D.M.A., Fairlamb,A.H., Fraunholz,M.J., Roos,D.S., Ralph,S.A., McPadden,G.I., Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C., Camucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and Barrett,B.		
TITLE	Genome sequence of the human malaria parasite Plasmodium falciparum		
JOURNAL	Nature 419 (6906), 498-511 (2002)		
PUBMED	12368864		
REFERENCES	2 (bases 1 to 271546)		
AUTHORS	Gardner,M.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-SEP-2002) The Institute for Genomic Research, 9712		

Medical Center Dr., Rockville, MD 20850, USA

FEATURES

source

1. 271546  
/organism="Plasmodium falciparum 3D7"  
/mol\_type="genomic DNA"  
/isolate="3D7"  
/db\_xref="taxon:36329"  
/chromosome="11"

repeat\_region  
128..152  
/rpt\_type=tandem  
complement(130..207)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
150..183  
/rpt\_type=tandem  
complement(313..373)  
/rpt\_type=tandem  
/rpt\_unit="(a)n"

repeat\_region  
313..343  
/rpt\_type=tandem  
complement(426..445)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
complement(485..504)  
/rpt\_type=tandem  
/rpt\_unit="(a)n"

repeat\_region  
complement(644..683)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
644..671  
/rpt\_type=tandem  
complement(786..836)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
799..836  
/rpt\_type=tandem  
complement(982..1091)  
/note="AT rich"

repeat\_region  
/rpt\_type=tandem  
complement(1002..1024)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
1148..1175  
/rpt\_type=tandem  
complement(1148..1175)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
complement(1245..1286)  
/rpt\_type=tandem  
/rpt\_unit="(a)n"

repeat\_region  
complement(1346..1412)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
1346..1373  
/rpt\_type=tandem  
complement(1380..1446)  
/rpt\_type=tandem  
1384..1433  
/rpt\_type=tandem  
complement(1412..1446)  
/rpt\_type=tandem  
/rpt\_unit="(a)n"

gene  
1534..2550  
/locus\_tag="PF11\_0457"

mRNA  
1534..2550  
/locus\_tag="PF11\_0457"

CDS  
1534..2550  
/locus\_tag="PF11\_0457"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="FAN36037.1"  
/db\_xref="GI:23496381"  
/translation="MEKEKITYDRQLRMGVKAKNRMKSNVLVVGLSGINELCKRLI

LANGINITIIDNNVDEEDIEINIFELNHEMKYNSVPIPELKSINOLINIKSFIGHI  
NLDITNNIIIEKELMYKDNIEIIQEBKINDCSYDIDYITNTVCVISCEDYPLKLT  
INIEIENKINIFPAMNGKAYAFLEPFGKHIIIESYYKIKNDNKGNDNSIEVOYC  
TISHPLKVPENLNDKTDNDIIYYFALILYEODKKMKONKOIDEQELNFTNNLANK  
TNTONPFLCKTYINKNSPSCSINGVTSQIRKFIHQHESINFCVFDNNQIVCT  
SMIX"

repeat\_region  
complement(1776..1830)  
/note="AT rich"

repeat\_region  
2544..2577  
/rpt\_type=tandem  
complement(2555..2620)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
2575..2620  
/rpt\_type=tandem  
complement(2664..2702)  
/note="AT rich"

repeat\_region  
2720..2748  
/rpt\_type=tandem  
/rpt\_unit="(a)n"

repeat\_region  
2750..2775  
/rpt\_type=tandem  
complement(2750..2775)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
2843..2881  
/rpt\_type=tandem  
complement(2843..2881)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
complement(2929..2962)  
/note="AT rich"

repeat\_region  
2990..3019  
/rpt\_type=tandem  
complement(2990..3019)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
complement(3027..3056)  
/note="AT rich"

repeat\_region  
complement(3113..3137)  
/note="AT rich"

repeat\_region  
complement(3160..3182)  
/note="AT rich"

repeat\_region  
complement(3221..3263)  
/note="AT rich"

repeat\_region  
3350..3401  
/rpt\_type=tandem  
complement(3350..3401)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
complement(3414..3506)  
/note="AT rich"

repeat\_region  
3503..3555  
/rpt\_type=tandem  
complement(3507..3576)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
complement(3578..3601)  
/note="AT rich"

repeat\_region  
complement(3667..3748)  
/rpt\_type=tandem  
/rpt\_unit="(ta)n"

repeat\_region  
3714..3748  
/rpt\_type=tandem

```
repeat_region complement (3760.. 3796)
                /note="AT_rich"
                /rpt_type=tandem
repeat_region complement (3873.. 3924)
                /rpt_type=tandem
                /rpt_unit="(ca)n"
repeat_region complement (3942.. 3967)
                /note="AT_rich"
                /rpt_type=tandem
repeat_region complement (4068.. 4164)
                /note="AT_rich"
                /rpt_type=tandem
repeat_region complement (4087.. 4119)
                /rpt_type=tandem
                /rpt_unit="(caaa)n"
                /rpt_type=tandem
                /rpt_unit="(caaa)n"
gene <4155.. >5108
mrna <4155.. >5108
CDS /locus_tag="PF11_0458"
    /locus_tag="PF11_0458"
    /locus_tag="PF11_0458"
    /product="hypothetical protein"
    /protein_id="AA036038.1"
    /translation="MNSKENDSVSKLSEKNNIKLIVKPEVSNKMLQYNNKDIYGL
    LQNNNNNSDITELYQKDDSDYKLRCKNNVTYIILKQVYIKLSTNSKSVSN
```

Query Match 1.2%; Score 121; DB 2; Length 271546;

Best Local Similarity 42.8%; Pred. No. 2.6e-07;

Matches 1766; Conservative 0; Mismatches 2265; Indels 96; Gaps 19;

```
QY 5120 ATCGTGTTCTGATGAAAGGAAAGCTTTAGTATTCGATTTGGAAGAGTAAGTT 5179
Db 188710 ATACTGAATTTGTTGTAAGAAAGATCAAGCTACTGAATTTGTTGTAAGAAAGATCG 188769
QY 5180 CTTTCAATGATGTAAGAAAGATCAAAACAAATAATCAATTCGCGAAATG---CAA 5236
Db 188770 CTACTGAAGTAGTTGTGTAAGAAAGATCAAGCTACTGAAGAGTTGTGTAAGAAAGATCG 188829
QY 5237 ATGGAAGAAAGAGCGCTGAGAGTTGAGCAACAGTTCATACAAATTTGGAAGAAACAT 5296
Db 188830 TTTCTGAAGAAATGTTAGAAAGAAAGATCAAGCTACTGAAGAGTTGTGTAAGAAAGAT 188889
QY 5297 CAGTTATAGCTATTGTAAGAAACAGTAATATTACAACGCGAATGATCAAGATAGAAAA 5356
Db 188890 TATCTAGTATATGTTCAAAAGTCAAAAGAGTATTTGAAAACGTAGAGAAATATATT 188949
QY 5357 ATATCAATGTGACTGCAAAAGATTATATGATGACCAATACTATAGCAGTCGAGATTGAG 5416
Db 188950 CTGTAAAAACAGCTAAGATGATCTATGATGAAAAAATACCATTAGAGAAATCATCAT 189009
QY 5417 GAGCAAAAGAGAGCTCTGTCAGAGAGCTTCTGCAAGTACTACTGATTAAGACGTTT 5476
Db 189010 TTGTTGATATAGTCTCTTTAAAGGACAAAGACCTACAGATATGATATAGTCGAAAGCG 189069
QY 5477 CTTTCATGTTGATCAACATGATATTGACAAAGATTAGAGAAAGAAATTAATGAAATA 5536
Db 189070 TAAATTCGAGAAATATTTAATTAAGACATATCTCATTTAGAGAAAGAAATTTGAAAGAT 189129
QY 5537 AGGAAAAAGCAATGTTAATGTTCTAGCTGAAATAACGA--GTCAAGTGTCACAAAATG 5593
Db 189130 TACCAACAGAAATATATTACTAGCTGATATACACACTAAAGGTGAAAAAGAAAAAT 189189
QY 5594 CGAGATGCTTTCCGAGAGCAAGTCAGACAGCTGCAATAGAGCTGAGAGTGAAGCTTAA 5653
Db 189190 ATATCTAATATATGCAAAAAATTAATGAAGAGTTAAAAAGAAAAATTTCAAGAGAAA 189249
QY 5654 AAATTCACAAAAATCTCTGACATATAAAAATAGTACTCAAAATGTACGAAATGCTT 5713
Db 189250 AAATTCACAAAAATCTCTGACATATAAAAATAGTACTCAAAATGTACGAAATGCTT 5713
QY 189250 AAATTCACAAAAATCTCTGACATATAAAAATAGTACTCAAAATGTACGAAATGCTT 189309
QY 5714 TGGTAAAAAGCAAACTCATCTCATTTAAAAACAATTGGAATTGAGCTGAGTTGAG 5773
```

```
Db 189310 CATCATTTGCTCAGAGAGTTGAACAAAAAGATCAGATATCAATGAATTTGAA---GAAA 189366
QY 5774 CTGAGAGAGCTGAGATGACAGGTTCTGAGACAGTAATAGATTGTAATATATACATAG 5833
Db 189367 CTAAATTTGCTTAATGAAAAAATTTATACAGAGATATCTGTTACTGAAAAACGTTGTTGAAA 189426
QY 5834 CAGAAATTAATCATGCAAAAAATCACTGCGAAAGGAAATGTCGAGTTATTACAGAGTCTG 5893
Db 189427 AAGAAAGATTTAGATATCTGAAAGAGCTTATGAAAGAGATGAATCAATTTACTGAAAGAAATAG 189486
QY 5894 ATGCGGTAATTTGCTAATTTATGACAGAACAGTGTGAGAGTGCCCGTCAGCAATAGAG 5953
Db 189487 TAGAAGAAAGAGTATCATCTGTGTGAAAGAAATAGTAGAAGAAAGAAATCATCTAGTGAAG 189546
QY 5954 CCTCAACCAAGTGAATGAAATTTACAGAGATCTACAAAGCATATGTAAAGATTCTACAG 6013
Db 189547 AAATGTAGAGAGAAAGAAATCATCTAGTGAAGAAATAGTAGAAGAAAGAAATCATCTA 189606
QY 6014 TGATTTGCTAAAGAGAAACAGATGATTTATTTACT-----ACTCAAGGCAAGTAGATA 6067
Db 189607 GTGAAGAAATAGTAGAGAGAGAAAGATCATCTAGTAGAAGAAATAGTAGAGAGAGAAAT 189666
QY 6068 AAGTGTGATTAAGTATTTCAAAAATCTTAATATTAAAGAGACTTATCAAAAAAGAA 6127
Db 189667 CATCTAGAGAAAGAAATTAAGAAAGAGTCAAGTACTGAAAGAGCTTGAAGAAAGAG 189726
QY 6128 AAATAGTAATTAAGAAAGATTTGTTACCAATAGTTGAGTCTACTCATCTTTAAATCTT 6187
Db 189727 GATCGGTTACGAGAGAAATTTGTAGAGAAAGATTCATCTACTGAAAGAGTAAAGATA 189786
QY 6188 TATTGGCAAAATGCGCGTGTTCAGAGCAAGCCGAGTGGCAAGAACTGTAATATCAACA 6247
Db 189787 TTGATCAGTAGTAGCAAGAGATATTAGAAAGAAAGAGATTCGCTAGTGAAGAAATTTGAG 189846
QY 6248 AAGTTTATGGA---GAAACAGAGCTCTTTGAGAAATTTCTATATTAATGCAAAACATT 6304
Db 189847 GACCAAGAGATCTGTTATTGAAAGAAATTTGAAAGAAAGAAATCACTCTGAAAAAGTTC 189906
QY 6305 ATTCGTAAATTAAGAGATTTACAGAAATTCATCGAGTATAGATTCTGTGTGTTG 6364
Db 189907 TAGAAGATGTTGATCAAAATTTGAAAGAAATAGTACAGAAAGAGTCTGTGTCTCAAG 189966
QY 6365 G---TGAAATGTAGAGTAGAGAGCTTCTGATACCAATATTTAAAAAGAAATACCA 6421
Db 189967 AAATTTGATCAGAAAGAGATTAATCACTACTGAAAAACATGATGAGTAGATCAACTA 190026
QY 6422 AGACAGAGTTGAAAACTA-----CAATGTCTGATGAGAGTTTCGAGAAAGAGCTG 6475
Db 190027 CTGAAGAAATTTGTAAGAAAGAGTTGATCAGTTCTGAAAGAAATTAATGTAAGAAAGTGT 190086
QY 6476 AAATTAAGCAGATTTCTAAGCAAGAAATTTCTCTTTTGGAGTGGAGTCGACAGCCG 6535
Db 190087 CAGCTTCTGAAGAAATAGTAGAAGAAAGATCAGTTACTGAAAGAGTTGTAAGAGGAAA 190146
QY 6536 GGGTAGAGCCGAGATGTCAGAGAACGTTTCGTAATCAATTTGACAGAAAGACGAG 6595
Db 190147 AATTAAATTAAGAGATTGTTGTAAGCAAGATCGGTTACTGAAAGAAATAGTACAAAAAGAG 190206
QY 6596 TAGATGGAAGAGCAAGATTTTGTAAAA-----AAAGTGAATTTACAGCA 6646
Db 190207 TATCAGATGCTGAAGAAAGATTTTATGAGCAAGAGATCAATGATTAAGAAATATTAGAA 190266
QY 6647 AACGTTATAGTTCTGTGCAATTTGAAATTTGCGCAGTGGAGTGGCTCAAAAGAGCTG 6706
Db 190267 AAGAAATCAATTTGTTAAAGAAATTTGTTGACCGAGAGATCTGTTACTGAAAGAAATAGTAG 190326
QY 6707 GAATTTGAGCA---GCAATGCGAGTTACCAAGATGAATCAAAACAGAGACAGAGTGA 6763
Db 190327 ATCAGAGATTCATTTGCTGAAGAGTCAAGAAAGAAAGAAATTTAGTTACAGAGAGCTGTAC 190386
QY 6764 AAATTTCTAAATTTATGACTCGAAACAGATTAGATGTATATGACAGAAATAGAGATTAAT 6823
```

Db 190387 AGTATGAAGGATCTGTATCTGATGAAGAAATAAAAAGAAAGAAATCGATTAAGTGAAG 190446  
QY 6824 CAGGACTGGAATCGGTTCAAGCCGAGCTGGAATCTTGACCCGAGTATTCGAGTGG 6883  
Db 190447 CAATGAGAAATCAACCTTTTGGCAAAATATATGAAAGAAAAGAACCAATATCTGACGAA 190506  
QY 6884 TTTCTGCAATTAATTTGCAAAATAGGTAGAAACGATTCGATCATAGTACTTTACCT 6943  
Db 190507 TTGTAAAAGAAAGAAATTAATGATTAAGTGAAGAAATGAGTAAATGATTAAGTACTGACG 190566  
QY 6944 CTTCTAGTATGTAATGTAAAGCTCTTAATAAATTTGCAATTCCTTGACGCGGTG 7003  
Db 190567 AGGTATGAGAAAGAAATTAATGTTAATGAACAATTTAGTGAAGAAAGATCTGTTA 190626  
QY 7004 GAGGAGCCGAGCTCTTGACAGATTTACCGAGTGGTTTCTGTAACTATTAATAGTT 7063  
Db 190627 CTGAAAAACAGTAGAAGTTGAACGATCAGTACTGAGATCTGTGTAAGAAAGCAAT 190686  
QY 7064 CTGTATAGCTCGAGTTCACAAATTAATCTGATTTGATCTCGTAGAGAAAAGTAAATG 7123  
Db 190687 CGGTATCTGAAAAAGTTTCACTGATGCAAGGATCAATCTAATC--AAATTAAGATG 190743  
QY 7124 TAAAGCAAAAAGAGAAAAAATATTAAGCAAAACAGCAAGCAATGCAAGATCGAGAG 7183  
Db 190744 AATCAAGTCTGAAAGAAATAGTAGAGAAAGAGTATCTGTGATGCAAAATTAATGAG 190803  
QY 7184 CAGCAATCGAGCCCAATGCTTGTGAATTAATTTTGGAAACAGCTGTAGAAATGAAAA 7243  
Db 190804 AAGAGTATCTGTGATGAAGTTGTAGAAAGAAAGATCAGTATTTGAAAGATGTAG 190863  
QY 7244 ATTTGAGAGAAAGAAAGAAAGTTTAAAACTTTAGACAGATTAACAAAGAAAG 7303  
Db 190864 AAGAGAAAGAAATCAAGTTCTGAAAGAAATTTGGAAAGAACTATCAGGTTCTGAAAG 190923  
QY 7304 ATAAAAAATTAATGATGCTAGCAAAAAAATCTTAATCAGAGGTAATTTCTACAGAG 7363  
Db 190924 TTCTAGAAACGAAATGGGTTACTGATGCTTTATGGGACAAAGAGATCTGTTATGAG 190983  
QY 7364 ATACTTCTGTAAAAGCGATAGAGAGATATCTCAGGAGAAAGAAATTTAAAGCCATGTGA 7423  
Db 190984 AATTAAGAAATATAGTAGATGAGAGATCGATTAACGAAAGAAATTTGAAAGATGAT 191043  
QY 7424 AGACTTCTGATTAATTTGAAAAAATGTAGATTTCAACAGAGGCAAGAAATATATCA 7483  
Db 191044 CAGCTAATGAAAAAATAGTAGAGAAAGCAATCTGTGTGAAAGATTTTAGAAAAAG 191103  
QY 7484 CTTCTAAGTGTGGTGGGAACTGCAAGCTTGCTCCGATCAGGAACAGTGGCAGTTA 7543  
Db 191104 AAGGATTTGTTATTAAGAAATATATAGAAAGGTCGGTAAATTGAACAAGTTGAAGAT 191163  
QY 7544 CAAATATTAAGAAATTCGAGAGTTACTGTGAAAAATCTTTTGTGAAGCAGCTGAAA 7603  
Db 191164 CA-----AAACAGTTCTGAAAAAATCAGAAAGATCATCAGTATTAAGAAAGTTAAAG 191217  
QY 7604 AAGTAATGTAGATCGGATATTAACAGAAATGTTGCTTTAACAGCAATCAAGTCTCG 7663  
Db 191218 AAGTAAAGAAAGAAATCATTAATAGTAAAAAATAGTAAAAAAGAAATCTGTACTG 191277  
QY 7664 TAGGAGCTTGGGAATAGAGAGCTGCCATAGCAAAATTAATTTCTAATGAAAGATCAAA 7723  
Db 191278 AAAAAATTTGACAAAGAAAGATCTAATCTAATAAAAAATAGTAAAGATGTATCA--- 191333  
QY 7724 TCAGTATTAAGAAATTTCTAAGCTATTAAGAAAAAATATATGATTTATTTGAAGATTAAT 7783  
Db 191334 --CCTACTGAAGATTTTGTGAACAACAAACATTAAGTACTGAAAAAGTTATGAACAAG 191391  
QY 7784 CGGAATTTGAGAGCGGAAAGCAAAAGATTAACCTAGAGCGGTAGCTCCGAGCCATTA 7843  
Db 191392 AAGGATCAAAACACTGAAGTTGACAGAAAGCCTTGAAGAAAAAGAAATCCGCTTCGATGAC 191451  
QY 7844 TCTCAAAAGCAAGAAATGAATTAATCAGAGGTGAATTTGAAGAGATATTTTCATG 7903  
Db 191452 ATGAACAAGAAAGTATCATCATTAATGCAAGTTACATATGAAAAAATCAGTTACTA 191511

QY 7904 AAGAA--AATAGATTAAGTACCCCTTCTAAAGAAATTTGAAGAAATCAATGTCA---- 7957  
Db 191512 AAGAAATGTGACGAAAGTGTCTGCTACTGAAGAAATAGTAGAAAGAAATGATCAAGTTA 191571  
QY 7958 -----AAGTGAAGAAAGAAAGAAAGAGTACTGCTGAATCTCAAGAGCTTCTGTAGAG 8011  
Db 191572 CGGAAGAGTGAACGAACAGGATCCGTTACTGAAGAAATATATGAAAGAACAGATTA 191631  
QY 8012 CAGTACGAGGGGACGAATTAATTTCCGAAGCAAAAGATCCGGAAGCTCTTAATTTGAAG 8071  
Db 191632 CTGAAGAGTGTGTAAGAAATGATCAAGTGAAGAAAGTGTGAAGATGATCAAGTTA 191691  
QY 8072 TTAGTACAAAATCCGGAAGAAATTTTTCATGCAATATGTGAATATGAAAGCAAC 8131  
Db 191692 TTCAAGAAAGTTGTAAGATGATGATTAATCTGAAGAAATAGTACAAAGAAATGATCAG 191751  
QY 8132 ATAAAAATGAAGTAAACAGCAAGTTCTTAAGCAAGTAAACAGTTCTGTATTTGGAGAGTGG 8191  
Db 191752 TTACTGAAGAAATATAGTAGAGAAAGAAAGATCAGTTAATGAAGAAAGTGAAGAAATGAT 191811  
QY 8192 GAGTCAAGGACAGAGCTACTGCTGACGTTAAACTATGTGTAGAAAGTTGAAAGAA 8251  
Db 191812 CAGTTTCTGAGAAAGTGTATGAACAGAAATATGTTACTGAAAGAGTGAAGAAAGAT 191871  
QY 8252 ATTTGTTCAAGCAAAATGATTTGAATGCAATTTCC--TAAAGTAAAGGTTTGAATGAAG 8308  
Db 191872 CAGTTGTTGAAGAAATTTGTAAGAGAGAAAGATCAGTTCTGGAAGAAAGTTGTAAGAAAG 191931  
QY 8309 ATTAAGTAACTGCTAAATCTTCTGTATATCAGAAATGAGAGAAATTTGCGAGCAG 8368  
Db 191932 AATATGTTATCTGAAGAAATGAAAGAAAGATCAAGTTGTTGAAGAAATTTGAGAGAG 191991  
QY 8369 GAGTAATATCTTCAAGCAAAAGTAAATCTGAATCCGTAGTTGTTAGCAAGCAAG 8428  
Db 191992 AAGGATCAGTTCTGAGAAAGTTGATGAACAGAAATATGTTACTGAAGAAAGTGAAGAG 192051  
QY 8428 A-----TAAAGAAATTAATGATTAACAAAAAATATTAATTTTCAAGATCAATGCTC 8479  
Db 192052 AAGGATCAGTTGTTGAAGAAATTTGTAAGAGAGAAAGATCAGTTCTGGAAGAAAGTTGATG 192111  
QY 8480 TTGCTTTAAATGATTAACAAAGAAATGAAAGCAATATAGAAATCTTTAGCCGTAGCCGCTGTC 8539  
Db 192112 AAACAGAAATGTATCTGAAGAAATGAAAGAAAGAAAGATCAAGTTGTTGAAGAAATTTGAG 192171  
QY 8540 ATGCAAGAAACAAACAAAGATTTAGAGATCAACAAAGTTAACTTTACACTGTAA 8599  
Db 192172 AGGAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAGAAAGATCAGTTCTGAAGTGA 192231  
QY 8600 ATGAGAGAAAGTAT-----CTCAACTTCGTCAAAAGCTTTGGCTAAAAATGAAA 8650  
Db 192232 TTGATGAACCTAAATTAATTAATGAAATGAAATGAAACAGCTCAATTTACTGAAGAG 192291  
QY 8651 ATTAATGAAGATTAAGAAAGAACTGAGAGAGCTTATGTCGAGCGGAAACAGCAGCCGTTG 8710  
Db 192292 TTGAAGAAAGATTAATCAAGTTAAAGATTAATAGAAATGATCAAGTTGCTGAAGCAG 192351  
QY 8711 AAAATTAATCAAGAGTACTACAGAGCAATGCTGTCAGAGAAATTTGGAATTTGAGAT 8770  
Db 192352 TTGAAGAAATCAAGATTAATTAATGATCAAGTTCCCAAGAA-----GAAAGAAACAGAAA 192405  
QY 8771 AATTGAAGAAAGATTCAGAGAGATTAATGATTTGTAAGATTCAGAGAGAGCAAGCAAG 8830  
Db 192406 AAGGATTTGTTATTTGAAAAAGTTGAAGAAACAGAGCCGTTACTGAAGAAATTTGTTCAAG 192465  
QY 8831 GAGGCTGTGCGGAAGAAATGATTTCTGTGAAAAATCAATTTTCAGGGGAAACAAAT 8890  
Db 192466 ATGGAATTAATTAACAGAAAGAAATTTGAAGAAATCTGAATCCGTTATGGGAAATTAATA 192525  
QY 8891 CATCAATGAAGTAAAGCAGAAATTTGGAACCGGAAGTGAATGTAGATGCTTTGA 8950  
Db 192526 ACAGAAATCAAGATCTGAGAAATTTTAGAAACAGAAATTTCTTACTGAAGAAAGTTGAG 192585









QY 6469 GAACTGAAATTAACAGCAGATCTTAAGCAAGAAATTCCTTTTGGAGTCGAGTCGCA 6528  
 Db 16873 GCATCAGCAGCTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCATTCACAGATTGGAGTACCT 16814  
 QY 6529 GCAGCCGGGGTAGGAGCCGAGAGTGACAGAAACCGTTCCGTAAATCAATTTTGCAGAAAG 6588  
 Db 16813 GTGGCAGAGAGATACAGATCAGCAGATTGTCAGTACGTGTGGCAGAGATAGCAGACGA 16754  
 QY 6589 ACGAAGTAGATGTGGAAGAACCAAAAGATTTTGGTAAAAAAAGCTGAGATTACAGCAAA 6648  
 Db 16753 GCAGTTGACAGTAGCTGTGGCAGAGAGTAGCAGCAGCAGCAGATTGCACTAGCTGTGCAGGA 16694  
 QY 6649 CGTTATAGTTCTGTTGCAATTGGAAATGCCGAGTCGAGTGGCTGCAAAAGAGCTGGA 6708  
 Db 16693 GTAGCAGCAGCAGCAGCTTGCAGTAGCTGTGGCAGAGAGTAGCAGCATTCAGCAGTTGCA 16634  
 QY 6709 ATTGGAGCAGCAGCTGGCAGTTACCAAAAGATGAAATCAAAACAGCAGCAAGAGTGAATAAT 6768  
 Db 16633 GTAGCTGTGGCAGAGAGAGTAGCAGCAGCAGCAGCAGATTGCCGTAGCTGTGCGCA 16574  
 QY 6769 TCTAAATTTATGACTCGAAACCAAGTTAGATGTAAATGAGAAAAATGAGATTAATCAGGT 6828  
 Db 16573 GCTGTTGACAGTAGCTGTGGCAGAGAGAGTAGCAGCAGCAGCAGATTGCCGTAGCTGTGCGCA 16514  
 QY 6829 ACTGGAATGGTTCAAGCCGAGCTGGAATTTCTTGACCCGAGATCTGAGAGTGTCT 6888  
 Db 16513 GAGGAGTTAGCAGCATCAGCAGTTGCGAGTAGCTGTGGCAGAGTAGCAGCATTCAGTAGTT 16454  
 QY 6889 GTCAATTAATTTGCAAAATAAGTAGAAGAACAGATATGATCATAGTCTTTACACTTCT 6948  
 Db 16453 GCAGTAGCTGTGGCAGAGAGAGTAGCAGCAGCAGCAGTTGCCGTAGCTGTGCGCAGAGGA 16394  
 QY 6949 ACTGATGTAAATGTAAAGCTCTTAATAAAATTTGCAATTCCTTGACCCGAGTGAAGA 7008  
 Db 16393 GTAGCAGCAGCTTGCACTAGCTGTGGCAGAGAGTAGCAGCATTCAGCAGTTGCACTAGTGTG 16292  
 QY 7009 GCCGAGCTCTTGACAGCAGTTACCGAGTGTGTTCTGTTAACTAATAAATAGTTCTGTG 7068  
 Db 16351 ATAGCAGTTGACAGTAGCTGTGGCAGAGAGTAGCAGCATTCAGCAGTTGCACTAGTGTG 16292  
 QY 7069 ATAGCTCGAGTTCAAAATACTCTGATTTGACTTCCTGACGAGAAAAAATTAATGTAAG 7128  
 Db 16291 GCAGAGAGAGTAGCAGCATCAGAGTTGAGTAGTCTGTGGCAGAGAGATACAGCATCA 16232  
 QY 7129 GCAAAAGAGAAAAAATTTAAAGCAAAACAGCAGCAATATGCAAGATCGAGAGAGAGA 7188  
 Db 16231 GCAGTTGACAGTAGCTGTGGCAGAGAGTAGCAGCATTCAGTAGTTGCACTAGTGTGCGAGA 16112  
 QY 7189 ATCGAGCCCAATGTCTTGTTAAATAATTTTGAACAGCTGTGAGAGATGAAAAATCT 7248  
 Db 16171 GAGTAGCAGCAGTTGCACTAGCTGTGGCAGAGAGTAGCAGCATTCAGTAGTGTG 16112  
 QY 7249 GAAAGAAAAAGAACAGAGTTTAAAACTTTAAGAGAGATTAAACAAAGATTA 7308  
 Db 16111 GCAGAGAGAGTAGCAGCATTCAGTAGCTGTGGCAGAGAGTAGCAGCATTCAGCAGTT 16052  
 QY 7309 AAAGTAATGATGCTACGAAAAAATCTTACAATCAGCAGGATTTCTACAGAAAGTACT 7368  
 Db 16051 GCAGTAGCTGTGGCAGAGAGAGTAGCAGCATTCAGCAGTTGCACTAGCTGTGCGAGAGTA 15992  
 QY 7369 TCTGTAAAAACCGATAGAGAGATCTACGAGAGAAAGAAATTAAGCCATTGTGAAGCT 7428  
 Db 15991 GCAGCAGTCGCGAGTTGCACTAGCTGTGGCAGAGAGTAGCAGCATTCGCGAGTTGAGCA 15932  
 QY 7429 TCTGATATTATGGAAAAAATGTAGATATTACAACAGAGAGCAAGAAATATCACTTCT 7488  
 Db 15931 GCTGTGGCAGAGAGATACAGCATTCGCGAGTTGCGAGCTGTGGCAGAGATAGCAGAGA 15872  
 QY 7489 ACTGTGTTTGGAGATGCAAGTCTTGCTCCGCAATCAGAAACAGTGGCAGTTCAAAAT 7548  
 Db 15871 TCGGCACTGTGCACTAGCTGTGGCAGAGAGAGTAGCAGCAGCAGATTGCACTAGTGTG 15812  
 QY 7549 ATTAAAAAGAAATTCGCGAGTTACTGTGAAAAATCTTTTGTGAAAGCAGCTGAAAAAGTA 7608

Db 15811 GCAGAGAGAGTAGCAGCAGTAGCAGTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCATCA 15752  
 QY 7609 AATGTAAATCGGAATTAATTAACAGAAATGTGCTTTAAACAGCATCAAGATCTCTTAGA 7668  
 Db 15751 GCAGTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCATTCAGCAGTTGCACTAGTGTGCA 15692  
 QY 7669 GCATTGGGAATAGAGCTGCTATGCAAAATTTAAATTTCTAATGAAAGATCAAAATATCAGT 7728  
 Db 15691 GGAAGAGCAGCAGCATTCAGTAGTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCAGACGA 15632  
 QY 7729 ATTAAAAATTTTAACTTAAGCTTAAGAAAAAATTTGATGTTTAAAGATTAATCGGA 7788  
 Db 15631 GTTGCACTAGCTGTGAGCAGAGTAGCAGCATCAGCAGATTGCGAGTGTGCTAGCAGAGA 15572  
 QY 7789 TTGAAGCGGAAGCAAAAGATTAACCTTAGAGAGCGGATAGCTGCCGAGCCATTATCTCA 7848  
 Db 15571 GTAGCAGCAGCAGCAGTTGCGAGTAGCTGTGAGAGAGAGTAGCAGCAGCAGCAGTAGCA 15512  
 QY 7849 AAAGCAAAAGAAATGAATGAATTCAGAGGTTGAAATTTGAAAGAGATATTTCATGAGAA 7908  
 Db 15511 GCAGCTGTAGCAGCAGAGTAGCAGCAGCAGCAGCAGAGTAGCAGAGAGAGTAGCAGCAGCA 15452  
 QY 7909 AATAGAGTACTAGCCCTTTCTAAAGAAATTTGAGAGAAATCAATGTCAAAGTGA 7968  
 Db 15451 GCAGTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCAGCAGCAGATTGCGAGTGTGCGCA 15392  
 QY 7969 GAAACAGAGTAGCTGCTGAATCTCAAGAGCTTGTAGAGAGCAGTGGCAGGCGAGGA 8028  
 Db 15391 GAGAGAGTAGCAGCATCAGCAGTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCATTCAGCA 15332  
 QY 8029 ATTATTTCCGAAGCAAAAGATGCCCCGAGCTTTATTTGAAAGTTAGTACAAATCCGGA 8088  
 Db 15331 GTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCATCAGTAGTTGCACTAGCTGTGCGAGA 15272  
 QY 8089 AGAAGTATTTTTCATGCAATATGTGAATGTGAATGTGAACAATGAAGTAAACA 8148  
 Db 15271 AGAGTAGCAGCAGCAGCAGTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCATTCAGCAGTT 15212  
 QY 8149 GCAGTTTCTAAAGCAGTAAACAGTTCTGTTATTTGAGAGAGAGTTGAGTCAACCAAGCAGAA 8208  
 Db 15211 GCAGTAGCTGTGGCAGAGAGTAGCAGCAGCAGCAGCAGTTGAGTAGGTTGGCAGAGAGAGTA 15152  
 QY 8209 GCTACTGCTGAGAGTAATACTATGTAGTAAGTTGAGAGAAATTTGTTCAACAAT 8268  
 Db 15151 GCAGCAGTTGCACTAGCTGTGGCAGAGAGAGTAGCAGCATTCAGTAGTGTGCGAGAGA 15092  
 QY 8269 CGATTGAATGCAATTTCTAAAGTAGAGCTTTGATGAATTAAGTAACTGCTAAATCT 8328  
 Db 15091 GCAGTAGCAGCAGCAGCAGCTTGCAGTAGCTGTGGCAGAGAGAGTAGCAGCAGCAGATT 15032  
 QY 8329 TCTGTAGTATCAGGAAATGAGAGAGAAATTCGCGAGAGAGAG 8371  
 Db 15031 GCAGTAGCTGTGGCAGAGTAGCAGCATTCGCGAGTTGCGAG 14989

RESULT 16  
 PFA929351/c PFA929351 347050 bp DNA 1linear INV 16-Apr-2005  
 LOCUS Plasmodium falciparum strain 3D7, chromosome 5, segment 1/4.  
 DEFINITION AL929351 AL844504  
 ACCESSION AL929351.1 GI:23504490  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 Plasmodium falciparum 3D7  
 Plasmodium falciparum 3D7  
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

## REFERENCE

1 Hall, N., Pain, A., Bartman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barton, A., Brooke, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corson, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,



Query Match	1.2%	Score 119.4	DB 2	Length 347050
Best Local Similarity	42.2%	Pred. No. 3.9e-07		
Matches	803	Conservative	0	Mismatches 1096; Indels 6; Gaps 2;
QY	7105	GTACGAGAAAAGTAAATGTAAACGCGAAAAAGAGAAATAATTAAGCAAAACAGACAGA	7164	
DB	45501	GGACAGAAATATATATACGAAAGAAATTAAGAGAAATTTAAAGAAATGATCTGAAAT	45442	
QY	7165	AATCGAGAAATCGAGAGACAGCAATCGAGCAATGCTTGTTAAATATATTTGGAC	7224	
DB	45441	AAAAGATAAGTGATAGGACAAAGAAATATATATCTGAGAGATTAATAAGAAATTTGAAAA	45382	
QY	7225	GCTGTAGAGATAGAAAAATTTCTGAAAGAAAAAGAACAGAAATTTAAAACTTTAGC	7284	
DB	45381	CAAGAGAAAAAGAAATTAAGAAAAATTTCTGAAATTTAAAGATATAGTAATTGGACAA	45322	
QY	7285	GAACTTAACAAGAACAGATTAATAAGTAATGATGCTACGAAAAAAATCTTACAATCA	7344	
DB	45321	GAACTTAATATAGAGAGATTAATAAGTAATTTAAAAAAAGTAAAGAAAAAGAAATTA	45262	
QY	7345	GCAGGATTTTCTACAGAGATATCTTCTGTAAAGCGAGTAAAGAGATATCTCAGAGAG	7404	
DB	45261	GAAATCATCTAGAAAGTAAGTAAGTAAGTATAGGACAAAGAAATATATAGTTGAAGAG	45202	
QY	7405	GGAATTAAGCATTTGAGAGATTTCTGATTTATTTGAAAAAATGTAGATTATACACA	7464	
DB	45201	AAAAGAGAAATTTGAAAAACAAGTAAAGAAAGAAATTTAAAGAAATGATCTGAAAGTAA	45142	
QY	7465	GAGGACAAAGATTAATATCATCTTCTACGTGGGTTTGGGAACTGACAGGTCTTGCCGA	7524	
DB	45141	GATTAAGTATAGACAAAGAGTATTAAGAGAGATTTATGAAGAGGTCCGAAAAAC	45082	
QY	7525	TCAGAAACGTGGACGTTACAAATATTTAAAGAAATTCGGAAGTTACTGTGAAAAATTC	7584	

Db	45081	AAAGATTAAGTGCACAAAACAGAAAAAGTAAAAAGATTAAAAAAGAACTAAAAA	45022
OY	7585	TTTGTGAAGAGCTGAAAAAGTAAATGTTAGATCGGATTTTCAGAAAATGTTGCTTTA	7644
Db	45021	GTTAAAAAAGAGTAAAAAAGAAATTAATAGAAATGAAAGAAAGATTAATGTGATAGCA	44965
OY	7645	ACACATATCAAGGTCTCTTAGAGCATTTGGGAATAGAGCTGCTATGACGAATTTAAAT	7704
Db	44961	AAAGAAATTAATGAAAGAAAGATGTTAATGAAAAAGATACCGCAACAAAAGTAAAGAGAT	44902
OY	7705	TCTAATGGAAGATCAAAATATCAGTATTTAAAAATTTAAGCTATTAGAAAAAATATTGAT	7764
Db	44901	GAACAAAAAAGAAAAAAGAAAGATTAAAAAAGAAAGAAAGATTAAAAAAGAAAGAA	44842
OY	7765	GTTATTGTAAGAAGTAAATCGGAATTGAGCGGAAACAAAAGATTTAACCCGTAGAGG	7824
Db	44841	GTTAAAGAAAAAGAAAGAGTAAAAAGAAAAAGAAAGTAAAAAAGAAAGAAAGTAAAA	44782
OY	7825	GTAGCTCCGAGCCATTATCTCAAAAGCAAAAGATGAATGAATTCAGAGTTGAAAT	7884
Db	44781	GAAAAAAGAAAGTAAAAAAGAAAAAAGAAAGTAAAAAAGAAAGATACCGAAAGCAAGAT	44722
OY	7885	GAGAAAGTATTTTCAAATGAGAAATAGATACAGCCCTTTCTAAGGAATTGGAAGA	7944
Db	44721	AAAGAGATTAAGACAAAGAAAAAGAAAAAGAAAGTAAAAAGATTAAAAAAGAAATACC	44662
OY	7945	GAAATCAATGTCAAAGTCGAAAAAAGAAACAGATGACGTGTGAATCTCAAGAGCTTCT	8004
Db	44661	GAAAACAAAGATTAAGTATGAGACAAAGATTAATATAGAAATAAAAAAAGAAAGTT	44602
OY	8005	GTAGAGCAGTAGACAGGAGCAGAAATTAATTTCCGAACAAAAGATCCCGAAAGCTTTAT	8064
Db	44601	AAAAAAGAGTTAAAAAAGAAATTAATTAATTAATGAACAAAGTAAATGTGATAGTACAA	44542
OY	8065	TTGAAAGTTAGTACAAAATCCGAGAAAGTATTTTCATGACAGATATGTGAATATGAA	8124
Db	44541	G---AAATATGAACGAAGATGTTAAAGAAAAAGATACCGCAAAACAAGATTAAGGTGATA	44485
OY	8125	GCACACATTAATTAAGAAAGTAAACGACGTTCTTAAGACGTAAACAGGTTCTGATTTGGGA	8184
Db	44484	GAACAAAAAAGAAAAAAGAAAGTAAAAAAGAAAAAAGAAAGTTAAAGAAAAAAGAA	44425
OY	8185	GGAGTTGAGTCAACAAAGCAGAAAGCTACTGCTGCAGGTAAAACTATGTGTAAGATTGAG	8244
Db	44424	GTAAAAAGAAAAAAGAAAGTAAAAAAGAAAAAAGAAAGTAAAAAAGAAAGTAAAA	44365
OY	8245	GAAGAAATTTGTTCAAGAACAAATCGATTGAATGCAATTTCTTAAGTAGAAGTTTGAT	8304
Db	44364	GAAAAAAGATCCGAAGCAAAAGTAAATATGTGATAGTAAAGAAATTAATGAAAGAT	44305
OY	8305	GAAATTAAGTAACTGCTAAATCTTCTGTAGTATCAAGAAATGAGAGAAATTGCCGA	8364
Db	44304	AACAAAAAAGATCCGAAGCAAAAGTAAATATGATGAAAAAGAAATTAATATGAAAGAA	44245
OY	8365	G---CAGAGTGAATCTTCTACAGACAAAGTAAATACGGAATCCGTAGTTGTTAAGA	8421
Db	44244	GTAAGAAAGAAAGTAAAAAAGAAAGTAAACAAAGAAAGTTAACAAAAGATTAACAGAGA	44185
OY	8422	AAAGCAATTAATGAATAATATGATTACAAAAAATATATTTCAAGAGTCAATGCTCTT	8481
Db	44184	AATGAAAAAAATGAAAAAAGATGTGATAGAAACAAGAAATGTGTAAGCAAGAAAGTTAAC	44122
OY	8482	GCTTTAAATGATACAAAGAAATGAGCGAATATGAATCTTTAGCGGTAGCCGCTGTGAT	8541
Db	44124	GAAAAAAGATCCAAAAACAACGATTAAGAAAGATGAAAAAAGAGTCAAAAAACAATAGAT	44065
OY	8542	GCAACAAGAACAAACAAGATTTAAGAGTCAACAAGTTAATCTTCAACTGTAAAT	8601
Db	44064	GATTTAAAAAAGAAAGAAAGTAAACAAGAAATCTGAAAGAAAGTCTTAAGAAAGAGTCT	44005
OY	8602	GAGAGAAACGTATCTCAACTTCTGTGCAAAAAGCTTTGGCTAAAAATGAATAATTATGAAAT	8661

Db		44704	GAAAGAAATCTGTAAGAAGAGTCTGAAGAAATCTGAAGAAAGTCTGAAGAAATCT	43945
Oy		8662	GTAAAAGCAACTGAGAGACCCCTTAGTCGAGCGGAAAACAGCACCCTTGAAATTATACA	8721
Dd		43944	GAAAGAAAGTCTGAAGAAATCTGAAGAAATCTGAAGAAAGTCTGAAGAAATCT	43885
Oy		8722	AAGGTACTACAGAGCATTTGGTTCGAGAAATTGGGAAATTGGAATTAATTGAAG	8781
Dd		43884	GAAAGAGTCTGAAGAAAGTCTGAAGAAAGTCTGAAGAAATCTGAAGAAATCT	43822
Oy		8782	ATTGCAAGAGATAATGATTTGTAAGATCACAGAACGAAACCAAAGAGCTTTGTC	8841
Dd		43824	GATGAGAAAAAATATACATCAGGTTGGTACATGAGAAATTTGTAAGAAAAAGAAA	43765
Oy		8842	GGAAGAATGTAATTTCTGTGAAAAATACATTTACAGGGAAACAAATATCATTTGA	8901
Dd		43764	TATATATATGAGAAATTAAGAAATATATTAAGAGAAACAGATGAAATATTTTGA	43705
Oy		8902	GATAAAGCCAGATTTGTTGGAACCGGAAGTGAATATGATGTTGGAATGAATTTGAT	8961
Dd		43704	GAGAAATATATATTTCAATCAAAAAGAACATPATCTTTGAAATCATTTCCCAATATG	43645
Oy		8962	GTAGTCTACAGGAAAAAGTGGGTGATGCTGGAATTTGGTAT	9006
Dd		43644	TTAAATGAAGATTTCCAGAGAAATTTCCACATATATTAAGTAT	43600
RESULT 17				
AF056936		5420 bp	DNA	linear
LOCUS				INV 10-MAY-2001
DEFINITION			Plasmodium falciparum mature parasite-infected erythrocyte surface antigen gene, complete cds.	
ACCESSION			AF056936	
VERSION			AF056936.1	GI:3044184
KEYWORDS				
SOURCE				
ORGANISM			Plasmodium falciparum (malaria parasite P. falciparum)	
REFERENCE			Plasmodium falciparum	
AUTHORS			Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.	
TITLE			1 (bases 1 to 5420)	
JOURNAL			Coppel,R.L.	
PUBMED			Repeat structures in a Plasmodium falciparum protein (MESA) that binds human erythrocyte protein 4.1	
AUTHORS			Mol. Biochem. Parasitol. 50 (2), 335-347 (1992)	
REFERENCE			1741020	
TITLE			2 (bases 1 to 5420)	
JOURNAL			Kun,J.F., Waller,K.L. and Coppel,R.L.	
PUBMED			Plasmodium falciparum: structural and functional domains of the mature-parasite-infected erythrocyte surface antigen	
AUTHORS			Exp. Parasitol. 91 (3), 258-267 (1999)	
TITLE			10072328	
REFERENCE			3 (bases 1 to 5420)	
AUTHORS			Kun,J.F.J. and Coppel,R.L.	
TITLE			Direct Submision	
JOURNAL			Submitted (01-APR-1998) Humanparasitology, Institute for Tropical Medicine, Wilhelmstrasse 27, Tübingen 72074, Germany	
FEATURES			Location/Qualifiers	
source			1..5420	
			/organism="Plasmodium falciparum"	
			/mol_type="genomic DNA"	
			/strain="Palo Alto"	
			/db_xref="caxon:5833"	
			join<1..150,281..>5116	
			/product="mature parasite-infected erythrocyte surface antigen"	
			join(1..150,281..5116)	
			/note="MESA; PFEMP2"	
			/codon_start=1	
			/product="mature parasite-infected erythrocyte surface antigen"	
			/protein_id="AAC13303.1"	
			/db_xref="gi:3044185"	
			/translation="MEVICRNLCYDKKNMMENEGNKVKRYNNSSLSKKYMKFCLTIICVFLDITYNCESTPYSSYSIMNNDYIILSETEPMSLEBIMEKTFEDHLYSIR	
CDS	mRNA			

	Query Match	1.2% ; Score 118.8 ; DB 2 ; Length 5420 ;
	Beech Local Similarity 41.9% ; Pred. No. 2.2e-06 ;	
	Matches 1240 ; Conservative 0 ; Mismatches 1672 ; Indels 48 ; Gaps 7 ;	
Oy	6204 TGGTTTCAGCAACAGCCGAGTGCAGGACTGTTAATATCAACAAGCTTTATGAGAACC 62633	
Db	1381 TGGTACAAAGAAAGTTCTGAAGAAACAGAAAGATGAACCGGAAGAAAATGAGAAAAA 14400	
Oy	6264 AGAAGCTCTTGTAATAATTCTATATTAATATGCAAAACATTATTCTGTAAATCAGAGAA 63233	
Db	1441 GGCAAGATTAATAAAAAAGTAAAAAAGAAAAAATCATTTTTTCAAATGTTAGCATGT 15000	
Oy	6324 TTACACGAATTCATTCGAGTAGTAGTCTGTTGGTGTGTCGGAATATGAGAGTAGG 63839	
Db	1501 TAAATTTCTTAATATTAATAATAATATGGAACCTGAGTAGAAGAAABAAACGTTGGTAGTAA 15666	
Oy	6384 AGCTTCTTCTGATACCA--ATATTATTAATAAAGAAATATCAACAAGACAGAGTTGGAAAAAC 64400	
Db	1561 AGATGATGCTTAATAAAGAAACATAAATTTTAAAGGAAGCTTAATCTGAAAAAATATGATTA 16220	
Oy	6441 TACAATGCTGATGAAGGTTTCGAGAGAAGAGCTGAAATTAACAGCAGATTCTAAGCAAG 65000	
Db	1621 TGAAAAAGAAAGATTAATATTAGAGAGAAGAGATTAAGAAGATGTTAAAGAAAAAGATGA 16880	
Oy	6501 AATTTCTCTTTGAGAGTCGAGATGCCAGACCCGGGGTAGAGACCGGAGTGCAGAGAAC 65666	
Db	1681 TGAACAGAAAGATTAAGATTATAGAGAGAAGAGATTAAGAAAGATTTAAAGAAAAAGATGA 17400	
Oy	6561 CGTTTCGTAATCAATTTGACAGAAAGACGGAAGTAGATGTGGAAGAGCAAGCAAGATTTT 66220	
Db	1741 TGAACAGAAAGATTAAGATTATAGAGAGAAGAGATTAAGAAAGATTTAAAGAAAAAGATGA 18000	
Oy	6621 GGTAAAAAAGCTGAGATTACAGCAAAACGTTATAGTTCTGTTGCAATTGGAAATGCCGC 66880	
Db	1801 TGGAAGAAAGATTAAGATTATAGAGATCAGAAAAACCAAAAAGAAATTAAGAAAAAGT 18666	
Oy	6661 AGTCGAGAGTGCACAAAAGAGAGCTGGAATTTGAGCACAGCAGTGGCAGTTACCAAAAGATGA 67400	
Db	1861 AGAAAAAAGAGATTAAAAAAGTGTAAAAAAGAAAAAAGTAAAAAAGAAATTAAGAAAAAGTGA 19200	
Oy	6741 ATCAAAACGAGACCAAGAGTGAATAATTTCAAAATTAATGACTGGAACAAGATTAGATGT 68000	
Db	1921 TAATGAAGATTAACGATTAAGTGAAGAGACCAAGAAATTAATTAATTT-----GAGAGAGT 19710	

Qy 6801 AATAGCAAAAATGAGATTAATACTGAGTACGGTTCAGCCGAGCTGGAATCT 6860  
Db 1972 AAAAGAGAAATTTAAAAACAAGTAGAAGATTAAAGAAATGTAAGTGAAGTAA 2031  
Qy 6861 TGCACCCGAGTATCTGAGTGGTTCTCTCAATATATGCAATTAAGTGAAGACGA 6920  
Db 2032 CGATTAAGTGAAGGAGCCAGAAATATATCTGAAGAGTAAAGAAATTTAAAAACA 2091  
Qy 6921 TATCGATCATAGTACTTTTCACTCTTCTGATGTAAATGTAAAGCTCTTAATAAAT 6980  
Db 2092 TGTAGAGAGGTATTTAAAGAAATATATCTGAAGTAAAGTAAAGTGAAGGAGCAGA 2151  
Qy 6981 TTGCAATTCCTTGACAGCCGCTGAGAGAGCCGAGCTTTCGACAGATTACCGAGTGT 7040  
Db 2152 AATATTAATCTGAGAGTAAAGAAAGAAATTTAAAAACATGTAGAGAAAGAAATTAAGA 2211  
Qy 7041 TTCTGTAACTATTAATAGTGTGTGTATAGCTGAGTTCACATTAATCTGATTTGAC 7100  
Db 2212 AATGATCTGAAAGTAAAGATTAAGTATGAGACAAAGAAATTAATTAAGTGAAGTAA 2271  
Qy 7101 TTCCTAGCAGAAAAAGTAAATGTAAACGCAAAAGAGAAAAAATATTAAACAAACAGC 7160  
Db 2272 AGAAGAAATTAAGAAAAATGATATGAAATTAAGTAAAGTATGAGACCAAGAAATGAT 2331  
Qy 7161 AGCAATGCAAGAAATCGAGAGAGCAATCGAGCCAAATGCTTGTGAATTAATTTGG 7220  
Db 2332 AACTGAGAGTAAAGAAAAAGAAATTTGAAAAACAAGAAAGAAAGAAATTAAGAAATAT 2391  
Qy 7221 AACACTGTAGAAATGAAAAATTTCTGAAGAAAAAGAACAGAAATTTAAAACTTT 7280  
Db 2392 TCTTAAATTAAGATATATGTAATTTGACAAAGATTAATTAAGAAAGTAAAAAAGT 2451  
Qy 7281 AGACGAAGTTAAACAAGAACAGATTAAGAAATTAATGATGCTAGCAAAAAAATCTTACA 7340  
Db 2452 AATTAAAAAAGATAGAAAAAGAAATTAAGAAATCTACT---GAAAGTAAAGATTA 2508  
Qy 7341 ATCAGCAGGTATTTCTACAGAAAGATCTTCTGTAAAGCGGATAGAGAGATCTCAGG 7400  
Db 2509 AGTGAATAGGACAAGAAATTAATGTTGAAAGATTAAGAAAGAAATTTGAAAAACAAGTGA 2568  
Qy 7401 AGAAGAAATTAAGCATTTGTGAAGACTTCTGATATTAATGAAAAAATGTAGATTAAC 7460  
Db 2569 AGAAGAAATTAAGAAAAATGATATCTGAAGTAAAGTAAAGTATGAGCAAGAAAGTAT 2628  
Qy 7461 AACAGAGCAAGAAATTAATATCACTTCTAGTGGTTGGAGACTGACAGCTTGTCTC 7520  
Db 2629 AAAAGAGATTTTAATGAAGAGGTCGCAAAAACAAGTAAAGTGAACAAAAACAGAAAA 2688  
Qy 7521 CGCATCGAGACAGTGGCAGTTACAAATATTAAGAAAAATTCGAGATTACTGTTGA 7580  
Db 2689 AGTAAAGAAATTTAAAAAGAAATTAAGAAAGAAATTTAAAAAGAGTAAAAAGAAA 2748  
Qy 7581 TTCTTTTGTGAAGCAGTGAAGAAAGTAAATGTTAGATCGGATATTAACAGAAATGTTGC 7640  
Db 2749 TAATAGAAATGAAGAAAGAAATGATATGTTAGAAAAAGAAATTAAGAAAGAAATGTTAA 2808  
Qy 7641 TTTAAACAGATATCAAGTCTCTGAGAGCAATTTGGATTAAGAGCTGCTATGACAAAT 7700  
Db 2809 TGAAGAAAGTACCGCAAAACAAGATTAAGAGATAGCAAGAAAAAGAAAAAGAAAGT 2868  
Qy 7701 AAATTTAAAGAGATCAATATATCAATTAATAATTTCTAAAGCTATTAGAAAAAATAT 7760  
Db 2869 TAAAGAAAAAGAA-----GTTAAAGAAAAAGAAAGTAAAAAGAAAAAGA 2916  
Qy 7761 TGAATGTTATTTAAAGATTAATCGAATTTGAGAGCGAGCAAGAAAGATTAAACGTAG 7820  
Db 2917 AAGAGTAAAGAAAAAGAAAGTAAAGAAAAAGAAAGTAAAGAAAGAAAGAAAGT 2976  
Qy 7821 AGCGGTAGCTCCGAGCAATTAATCAAGCAAAAGCAAGTGAATGAATTCAGAGTTGA 7880  
Db 2977 AAAAGAAAAAGAAAGTAAAAAGAAAGAAAGAAAGTAAAAAGAAAGTAAAAAGCA 3036  
Qy 7881 AATGAGAGATATTTCAATGAAGAAATAGAGTAACTAGCCCTTCTAAAGAAATGG 7940

Db 3037 AGATTAAGAGATAGAACAAAGAAAAAGAAAAAGAAAGTAAAGAAAGTAAAAAGAAAAAGA 3096  
Qy 7941 AAGAGAAATCAATGTCAGATGAGAAAAAGAAAAACAAGTGAATGCTGATCAAGAGC 8000  
Db 3097 TACCGAAAAACAAGTAAAGTATGACAGAAATTAATTAAGAAATTAAGAAATTAAGAA 3156  
Qy 8001 TTCTGTAGAGCAGTAGCAGGAGGAGAAATTAATTTCCGAAACCAAAAGATGCGGAGCTC 8060  
Db 3157 AGTTAAAAAAGAGTAAAAAAGAAATTAATTAAGAAATTAAGAAATTAAGTATGATAGT 3216  
Qy 8061 TTAATTTGAAGTATGACAAAAATCCGAGAAAGTATTTTCTAGTCAATTAATGTAAATAT 8120  
Db 3217 ACAAGAAATTAATGAACGAGATGTTACGAAAGATCCGCAAAACAAGATTAAGTATGA 3276  
Qy 8121 GGAAGCAACAKTA-----AAATGAAGTAAACGCAATTTCTAAGCAGTAAACAGTTTC 8174  
Db 3277 ACAAGAAAAAGAAAAAGAAAGTAAAGAAAAAGAAAGTAAAGAAAAAGAAAGT 3336  
Qy 8175 TGTATTTGGAGAGTGGAGTCAACCAAGCAGAAAGCTACTGCTGAGGTAAAACTATGAT 8234  
Db 3337 AAAAGAAAAAGAAAGTAAAGAAAGAAAGAAAGTAAAGAAAGTAAAGTAAAGCA 3396  
Qy 8235 AGAAGTTGAGAGAGAAATTTGTTCAACAATCGATTTGAATTCATTTCTAAAGTGA 8294  
Db 3397 AGATTAAGAGATAGAACAAAGAAAAAGAAAGAAAGTAAAGAAAGTAAAGAAAGAAAGA 3456  
Qy 8295 AGTTTGAATGAAGTAAAGTAACTGTAATCTTCTGATAGTACG-----GAAATGG 8348  
Db 3457 TACCGAAAAACAAGTAAAGTATGACAGAAATTAATTAAGAAATTAAGAAATTAAGAA 3516  
Qy 8349 AGAAGAAATTCGAGAGAGAGTGAATCTTCTACAGCAACAAGTAAATCTGAATCCGT 8408  
Db 3517 AGTTAAAAAAGAGTAAAAAAGAAATTAATTAAGAAATTAAGAAATTAAGTATGATAGT 3576  
Qy 8409 AGTTCTTTACGAAAGCAAGATTAAGAAATTAATGATTAACACAAAAAATATATTTTCA 8468  
Db 3577 ACAAGAAATTAATGAACGAAGATGTTAACGAAAGAAATCCGCAAAACAAGATTAAGTAT 3636  
Qy 8469 AGTCAATGCTGCTGTTAAATGATCAAAAGATGAACGAATATAGAAATCTTTAGCCGT 8528  
Db 3637 AGAACAAAAAAGAAAAAGAAAGTAAAGAAAGAAAGAAAGTAAAGAAAGAAAGAAAGA 3696  
Qy 8529 AGCCGTGTGATGACACAAAGAACAAACAAGCATTTACGAGATCAACAAGTTAACTTC 8588  
Db 3697 AGTAAAGAAAAAGAAAGTAAAGAAAGAAAGTAAAGTAAAGCAAGAAAGTAAAGATGA 3756  
Qy 8589 TACAACTGTAAATGAGAGAAACGTATCTCAACTTCGTCAAAAGCTTTGGCTTAATAATGA 8648  
Db 3757 ACAGAAAAAAGAAAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGTAAAGTAAAGTAAAG 3816  
Qy 8649 AAATTAAGAAATGTAAGAAAGTAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 8708  
Db 3817 TAAAGTATGAGACAAAGAAATTAATTAAGAAATTAAGAAAGTAAAGTAAAGTAAAGTAAAG 3876  
Qy 8709 TGAATATTAACAAGATTAACA-----GAGCATTTGTTGACAGAAATTTGGA 8759  
Db 3877 AAAAAAAGAAATTAATTAAGTAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3936  
Qy 8760 AATTGAGATTAATTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 8819  
Db 3937 CGAAGATGTTAAACGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 3996  
Qy 8820 CGGAACCAAGAGAGTCTTTGCGGAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 8879  
Db 3997 AAAAGAAAGTAAAGAAAGTAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 4056  
Qy 8880 GGAACCAAAATCATCTGAAGATTAAGCAAGAAATTTGTTGAAACCGAAGTGTAAATGT 8939  
Db 4057 ACAGAGATTAATTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 4116  
Qy 8940 AGATGCTTTGAATGAATGATGATGATCTACAGAAAGAAAGTGTGCTATGTTGAT 8999

Db	4117	TAATATAAATGAAAAACAAGATAATGTAGTACACAAATAATGACGAAGATGTTAA	4176
Qy	9000	TGCTATTGGAATGTTGATGTAATATGATTAAGAAAAATGTAAGACCAAAATCGG	9059
Db	4177	CGAAAAAGATACCCGAAACCAAGATAAATGATAGGAAAAAGATAATATAGAGAAAGT	4236
Qy	9060	AAAGCATCTATTGTAGAAACACTCGAAAAACAATAATCAAGATTTAACAGGACAA	9119
Db	4237	AAAAGAAAGCTTAAAAAAGATTAACAAAGAGTTTAACTAAAGATTAACAGAGAAA	4296
Qy	9120	AGTAATATCTCTTGGAAG 9139	
Db	4297	TAGAAAAATGAAAAAG 4316	
RESULT 18			
LOCUS	A58933	5361 bp	DNA
DEFINITION	Sequence 2 from Patent WO9641877.		linear
ACCESSION	A58933		
VERSION	A58933.1		
KEYWORDS	GI:3714407		
SOURCE	unidentified		
ORGANISM	unclassified		
REFERENCE	1		
AUTHORS	Drullhe, P. and Daubersies, P.		
TITLE	MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES		
JOURNAL	Patent: WO 9641877-A 2 27-DEC-1996; PASTEUR INSTITUT (FR) AU 6309696 970109 Other publication FR 2735478 961220. Location/Qualifiers		
COMMENT			
FEATURES			
source	1..5361		
	/organism="unidentified"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32644"		
	1..5361		
	/note="unamed protein product"		
	/codon_start=1		
	/protein_id="CAA03511.1"		
	/db_xref="GI:3714408"		
	/translation="MTSNYKSNKTYNENNEQITTTFNRTNMPICKCHREKINK YFPLKILCTLLIWAQYDNNDSINKSKNTTYDKLKNLPNSLGSQYNGELAS BEVEKILDLLEEGLTLESVDNKLAEADIKENILSNEPEPNTIDMLNNGIG ONSEKQESVENVQVDELFNELNLSVDNGVGENILLESQVNDI FNSLVKSVQGE OOHNVBEKESVEENDEESVEENVEENVEENDDGSVASVESIASVDESIDSIE ENVAIPTVEIIVAPSVESVAPSVESVEENVEESVAENVEESVAENVEESVA VAENVEIIVAPTVEIIVAPTVEIIVAPSVESVAPSVESVEENVEESVAENVEESVA ENVEESVAENVEESVAENVEESVAENVEIIVAPTVEIIVAPTVEIIVAPTVEI VEESVEENVEESVAENVEESVAENVEESVAENVEESVAENVEESVAENVEESVA ESVAENVEIIVAPTVEIIVAPTVEIIVAPSVESVAPSVESVEENVEESVAENVE VAENVEESVAENVEESVAENVEIIVAPTVEIIVAPTVEIIVAPSVESVAPSVESVE ENVEESVAENVEESVAENVEESVAENVEIIVAPTVEIIVAPTVEIIVAPTVEI VEESVEENVEESVAENVEESVAENVEESVAENVEESVAENVEESVAENVEESVA EIVAPSVESVAPSVESVEENVEESVAENVEESVAENVEESVAENVEESVAENVE TILLENVEETTESVTTFSNILEIEIOENTITNDTIEKLEBELHENVLSALENTOELE EKKEVIDVEIENKVEEYATTLIETVPOAKESANTTIEPIENLENNVSENNVAE KLNBEIVFNTVLDEKVEETVEISGESLENEMDKAFPEIITDNKGIQENILTMFRSIE TSIVIQSEBKVDLNNVSSILDNINMEGLINKLENTISTEGVEITVEHQNVY VVVDVAMKQDFGLINBAGGLKEMFPNEDVFKSESQVITVEIIDEBOVKQEKET VSIIEMENIIVDLBEKEDLTDKMIIDAVESIEISDSKEBTESIKQEKQVSLKVE BEVQDDMESVEKVELKMBEELMKDAVEINDITSIKIEFTQELNVEADLIKDEI KKELEKALSDESKETIIDAQDTEKVEIBEHDTITTLBVEELKQVEEDKIKVEDL KQLEEDILKEVKEIKELBSEILEDYKELTITIEDILBEKKEIKQHPKPEPEABERIK DLEADILKEVSSLEVEEKLEVEHLEKVEHISGDATIKQLEEDLBEVDLGS IIDLMLGDMELGDMDESLEDVTTKLGBRVESLKDVLSSMLQDEBQMTREKQAQPK LEEDILKEVEKEPKKTKKVRPDIKQEPKDEIIVEMKDEDEEDVEEDIEEDIE EBDKVEDIDEDIDEDIGEDKQVILIVQEKRIEKYKAKKKLEKKVEGSGSLKH VDEVMKYQKIDKDEVDKESKALSKNDLNVKONODPFSKYNPKKKVAAPEI SAVAASVYVGFPTSLBSSCTTIASTYLLSKVDKTKNKNERPFYSVPIPKNL KYLQMKKEKFSKKNNVLEVTNKAKEKNQVYTKTEKTTTVYDANKPKKRRQK		
CDS			

ORIGIN	SK"		
Query Match	1.24; Score 116.2; DB 6; Length 5361;		
Best Local Similarity	41.74; Pred. No. 4.9e-06;		
Matches 1747; Conservative	0; Mismatches 2373; Indels 67; Gaps 14;		
Qy	5126	GTTCTGATGAAACGGAAGCTTTAGTATGATGATTTCTGAGATTGGAAGAGTAAGTTCTTCA	5185
Db	773	GTGTTGAAGAAAGTATGACTTCAAGTGTGATGAAAGATATGATTTCAAGTATGAAAGAA	832
Qy	5186	ATGTAGATGCAAAAAGATCAAAAACATTAATACAA---TTGCCGGAATGCCAAATGAG	5242
Db	833	ATGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAAAGTGTGTGAAGAGTGCTCCAA	892
Qy	5243	GAAGACCGCTGAGATTGAGACACAGTCTCATCAAAATTTGGAAGAAACAATAGTTA	5302
Db	893	GTGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGAAGTGAAGTGAAGAA	952
Qy	5303	TAGCTATGTAAAAAAGCTAAATTTACAGGGCGAATGATCAAGATGAAAAAATATCA	5362
Db	953	GTGTAGCTGAAAAAGTGTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGAAGAA	1012
Qy	5363	ATGTGACTGCAAAAAGATTATATCTATGACCAATCTATAGCACTCGAGTTGAGAGCA	5422
Db	1013	ATGTTGAAGAAATGCTAGCTCCAACTGTGAAGAAATGCTAGCTCCAACTGTTGAAGAA	1072
Qy	5423	AAGAGCCTCTGTGCAAGAGAGCTTGCAGAGTACTGATTAATAGACAGTTTCTTC	5482
Db	1073	TTGTAGCTCCAAAGTGTGTGAAGAGTGTGCTCCAAAGTGTGMAAGAAAGTGAAGAA	1132
Qy	5483	ATGTTGATCAACATATATGACAAAGATTTAGAGAGAAATTAATGAAATTAAGAA	5542
Db	1133	ATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA	1192
Qy	5543	AGGCAATGTTATGTTTCTAGCTGAAATATCAAGATCAAGTGTCACAATGCGACAGTGC	5602
Db	1193	GTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA	1252
Qy	5603	TTTCCGAGCAAGTGCACAGCTGCAGTAGAGAGCTGAGTACAGTTAATAATTAAC	5662
Db	1253	ATGTTGAAGAAATGCTAGCTCCAACTGTTGAAGAAATGCTAGCTCCAACTGTTGAAGAA	1312
Qy	5663	AAAAATCTTCTGCACATATTAATAATATGATCTCAAAATGTAGAAATGTTGTGTAATA	5722
Db	1313	TTGTAGCTCCAAAGTGTGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA	1372
Qy	5723	GC---AAATCTCATCTATTAATAACAATTGGAATTGAGCTGAGTTGAGCTGAG	5779
Db	1373	ATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA	1432
Qy	5780	GAGCTGAGTGAAGAGTCTGTAGCAGTGAATAGATTGTAAATTAATGATAGCAGAAAT	5839
Db	1433	GTGTAGCTGAATAATGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTGAAGAA	1492
Qy	5840	TAAATCATGCAAAAATCACTGCGAAGGAAATGTGGAGTTATTCAGAGCTCTGAGCGG	5899
Db	1493	ATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA	1552
Qy	5900	TAAATCTGAATTAATGACAGAAACAGTGTGAGTGGCCGTCAGCAATAGAGCTCA	5959
Db	1553	TGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAA	1612
Qy	5960	CCAGTGTGAATGAATTAACAGATCTCAAAAACATATGTAAAGATTTCAACAGTATTC	6019
Db	1613	GTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAA	1672
Qy	6020	CTAAGAAAGAAACAGTATATTAATTAATCACTCAAGCGCAAGTAGTAAGTGTAGATA	6079
Db	1673	GTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA	1732
Qy	6080	AAGTATTCAAAAATCTTAATTAATTAACGAAGACTTATCAAAAAAAGAAATTAAGTAATA	6139



QY 8268 TCGATTGATGCAATTTCTAAGTAGAAGGTTGGATGAGATTAAGTAATCTGTAATC 8327  
DB 3936 CAGAGTCGAAAAAGATCTGATTTAAAGATCTTGAGAGATTAATTTAAAG----- 3988  
QY 8328 TTCTGTAGTATGAGAAATGAGAGAAATGCGGAGCAGAGATGTAATCTTCTACAGC 8387  
DB 3989 --AAGTAAAGAAATCAAGAACTTGAAAGTAAATTTTGAAGATTAATAAGAAATTAA 4046  
QY 8388 ACAAGTAATCTGAATCCGATGCTTCTTACGAAAGCAAGATTATGAAATTAATGATTA 8447  
DB 4047 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAAGATCATTTTGAAA 4106  
QY 8448 CACAAAAATATTTTGAAGTCAATGCTCTGCTTAAATGATACAAAGATGAAGC 8507  
DB 4107 ATTCGAAAGAAAGCTGAAAGAAATAAAGATCTTGAGCAATATTAATAAGAGATATC 4166  
QY 8508 GAATATGAAATCTTTAGCGGTAGCCGTTGTCATGACAAAGAAACAAAGCAATTTAC 8567  
DB 4167 TTCATTAGAAAGTGAAGAAAGAAATTAAGAAAGTACAGAAATTAAGAAAGAGGT 4226  
QY 8568 GAGATCAACAAAGTAACTTCTACAACTGTAAATGAGAGAAAGTATCTCACTCTGC 8627  
DB 4227 AGAACTATTAATAGGTGATGCGCATTAAGAGGTTGAAAGAGATGATTTAGAA 4286  
QY 8628 AAAAGCTTGGCTAAATGAAATTAATGAAATGTAAGAACTGAGAGACCTTAGT 8687  
DB 4287 AATAGATGA-TTTTAAAGAGATATATGACATGTTAAAGGAGATATGA-ATTAG 4343  
QY 8688 CGAGCGGAAACGACGCGTTGAAATTAACAAAGATTAATCAAGAGATTAATGAGAGC 8747  
DB 4344 GGAATATGATTAAGAAAGTTAGAAATGTAACACAAACCTTGAGAAAGAGCTGAATC 4403  
QY 8748 AGGAAATTTGGAAATTTGAGATTAATTAAGAAAGATTAATGATTAATGATTAAG 8807  
DB 4404 CTTAAGAAAGATTTTATCTAGTCACTTAAGCATGATGAAGAAATGAAGAAACAGAA 4463  
QY 8808 AGTCAACGAGACGAAACCAAGAGAGTCTGTCGAGAAAGATGATTTCTGTGAAA 8867  
DB 4464 AAAAGCTCAAGACCTAGTTGAAAGATTAATTAAGAAAGAGTTAAAGAGAAC 4523  
QY 8868 TCAATTTTCAAGGGAACAAATATCATCTGAAGATTAAGCAGAAATTTGGAACCG 8927  
DB 4524 AAAAGAAAAATTAACAAAAAGAAAGTAAAGCTTGAATTAATGAATTAAGAACCAAA 4583  
QY 8928 AAGTGAATTTGATGCTTTGATGAATCTTAATGATTAATGATTAATGATTAATG 8987  
DB 4584 TGAATATGATGAAGTTGAATGAAGTGAATTAAGAAAGATGATGAAGATAT 4643  
QY 8988 CTATGCTGAATTTGATTTGAAATGTTGATTAATGATTAATGATTAATGATTAAT 9047  
DB 4644 AGAAGAGATATGAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAT 4703  
QY 9048 AGCAAAATCGAAGACATGCTATTGAGAACTATCGAAGAAACAAAGATATCAAGATT 9107  
DB 4704 AGGTGAAGCAAGATGAAGTATTAAGTTTAATATGTCACAAAGAGAAAGCAATGAAA 4163  
QY 9108 TACAGAGCAAAATTAATTTCTTGAAAAAGAGACGCTGACGCTGACGCTCAATATC 9167  
DB 4764 GGTAAAGCGAAAAAGAAAAATTAAGAAAAAGTTGAAGAGGTGTTAGTGTCTTA 4823  
QY 9168 GAATGATACATTTCCATGAGATGATTAATTAATTTGGCAAGACATATGATCTTC 9227  
DB 4824 AAAACACGTAGCAAGATTAAGAAATATGTTCAAAAAATGATTAAGAAAGCTTAAGA 4883  
QY 9228 TCAATTAATACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9274  
DB 4884 AGATCTTAAGCTTTAGATCAAAAAATGATTTACTTAATGTTTAA 4930

RESULT 19  
ARI31272 5361 bp DNA 1linear PAT 16-MAY-2001  
LOCUS  
DEFINITION Sequence 2 from patent US 6191270.

ACCESSION ARI31272  
VERSION ARI31272.1 GI:14119597  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 5361)  
AUTHORS Drullhe P. and Dauberties P.  
TITLE Malerial pre-erythrocytic stage polypeptide molecules  
JOURNAL Patent: US 6191270-A 2 20-FEB-2001;  
FEATURES  
source location/Qualifiers  
1..5361  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 1.2%; Score 116.2; DB 6; Length 5361;  
Best Local Similarity 41.7%; Pred. No. 4.9e-06;  
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;  
QY 5126 GTTCTGATGAACGGAAGCTTTAGTATGATTCGAGTTGAAGGAGTAAGTTCTTTCA 5185  
DB 773 GTGTTGAAGAAAGTATAGCTTCAAGTGTGATGAAGATTAATTAATTAAGAA 832  
QY 5186 ATGTAGATGCAAAAGATCAAAATCAATTAATCAAA--TTGCGGAAATGCAAAATGAG 5242  
DB 833 ATGTAGCTTCAAGCTTTGAAGAAATCGTACCTCAAGTCTTGAAGAAAGTGTGCTCA 892  
QY 5243 GAAAACGCGCTGAGATTGAGCAAGCTTCTCATACAAATATTGAAACAAATCAATCA 5302  
DB 893 GTGTTGAAGAAAGTATGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA 952  
QY 5303 TATGCTATTTGTAATAAAACAGTAAATTAACACGCGCAATTAATCAAGTATGAATA 5362  
DB 953 GTGTAGCTGAATATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAG 1012  
QY 5363 ATGTAGCTGAAGAAATTTATCTATGACCAATATATAGACAGTGTGAGTGTGAGAGCA 5422  
DB 1013 ATGTGAAGAAATCTGATGCTCAACCTGTTGAAGAAATGATGCTCAACTGTTGAAGAA 1072  
QY 5423 AAGGAGCTCTGTGCAAGAGCTTGTCAAGTACTAATCTTGAATAAGACAGTTCTTCTC 5482  
DB 1073 TTGTAGCTCAAGTGTGTAAGAAAGTGTGCTCAAGTGTGAAGAAAGTGTGAAGAA 1132  
QY 5483 ATGTGATCAAACTGATATTTGACAAAGATTTAGAGAAAGAAATTAATGAATAAGAA 5542  
DB 1133 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAA 1192  
QY 5543 AGGCAAAATGTTATGTTCTAGCTGAAGAAATATACAGTCAAGTGTCAAAATGCGACAGTGC 5602  
DB 1193 GTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAG 1252  
QY 5603 TTTCCGAGCAAGTGAAGAGCTGACAGTGAAGCTGAGTGAAGTGAATTAATTAATTAAC 5662  
DB 1253 ATGTGAAGAAATCGTATGCTCAACCTGTTGAAGAAATCGTATGCTCAACCTGTTGAAGAA 1312  
QY 5663 AAAATATCTTGTGACATTAATAAAATATGATCTCAAAATGTAAGAAATGCTTTGTA 5722  
DB 1313 TTGTAGCTCAAGTGTGTAAGAAAGTGTGCTCAAGTGTGAAGAAAGTGTGAAGAA 1372  
QY 5723 GC---AAATCTCATTCATTAATAAAACAATTGGAATTTGAGCTGAGTGTGAGCTGAG 5779  
DB 1373 ATGTGAAGAAAGTGTAGCTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAA 1432  
QY 5780 GAGCTGAGTGAACAGGTTCTGTAGCAGTGAATTAAGATTTGAATTAATTAATTAATTA 5839  
DB 1433 GTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAG 1492  
QY 5840 TAAATCATGCAAAATCACTGCGAAGGAAATGTCGAGTATTAACAGACTGTGATGCGG 5899  
DB 1493 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAA 1552  
QY 5900 TAATGCTATTAATGAAGAAAGTGTGTGAGTGTGCGGTGAGCAATTAAGAGCTCA 5959

Db 1553 TCGTGGCTCCAACTGTGAAAGAAATCGTAGCTCCAACTGTTGAAAGAAATGTAGCTCCAA 1612  
 Qy 5960 CCAAGTGTGAATGAATTTACAGAGATCTACAAAGCAATATGTAAGATTTCTACAGTATGG 6019  
 Db 1613 GTGTTGTGAAGAAAGTGGCTCCAACTGTTGAAAGAAAGTGTGAAGAAATGTGGAAGAA 1672  
 Qy 6020 CTAAAGAAAGAAAGATGATTTATTTACTCAAGGCGCAAGTAGATAAAGTGTAGATA 6079  
 Db 1673 GTGTAGCTGAAATGTGTAAGAAAGTGTAGTGAAGAAATGTGTAAGAAAGTGTAGCTGAAA 1732  
 Qy 6080 AAGTATTCAAAATCTTAATATTAAAGAACTTATCACAAAAGAAAAATAGTAATA 6139  
 Db 1733 ATGTGAAGAAAGTGTAGCT-----GAAAATGTGAAAGAAATCGTAGCTCCAACTGTG 1786  
 Qy 6140 AAAAGAGATTTGTACCAATAGTACAGTACTCATCTTAAATCTTTATTTGCAAAATG 6199  
 Db 1787 AAGAAATCGTACTCCAACTGTTGAAAGAAATGTGTAGCTCCAACTGTTGA---GAAAGTG 1843  
 Qy 6200 CCGCTGTTCAGAGACAAAGCCGAGAGTGGCAGAACTGTTAATATCAACAAGTTTATGAG 6259  
 Db 1844 TGGCTCCAAAGTGTGAAAGAAAGTGTGAAGAAATGTTGAAAGAAAGTGTAGCTGAAAATG 1903  
 Qy 6260 AAACAGAACTCTGTGAAAGAAATTCATATTTAAATGCAAAACATTTATCTGTAATACAG 6319  
 Db 1904 TTGAAGAAAGTGTAGCTGAAAGAAATGTGAAAGAAAGTGTAGCTGAAATGTGAAAGAAATCG 1963  
 Qy 6320 GAGATTTACAGAAATTCATCGAGATAGTAGTGTCTGTTGGTGTGGTGAATGTAGAG 6379  
 Db 1964 TAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTGTAGCTCCAAAGTG 2023  
 Qy 6380 TAGAGACTTCTTCTGATACCAATATTATTAAG---AAATACCAAGCAAGAGTTGAA 6436  
 Db 2024 TTGTGAAGAAAGTGTGGCTCCAAAGTGTGAAAGAAAGTGTAGAGAAATGTGAAAGAAAGTG 2083  
 Qy 6437 AAACATCAATGTCTGATGAAGGTTTCGGAAGAAAGCTGAATTTACAGACATTTCTAAC 6496  
 Db 2084 TAGCTGAAATGTGAAAGAAAGTGTAGCTGAAAGAAATGTGAAAGAAAGTGTAGCTGAAATG 2143  
 Qy 6497 AAGGAATTTCTCTTTTGGAGTCCGAGTCCGACAGCCGGGTAGAGCCGAGAGTGCGAG 6556  
 Db 2144 TTGAAGAAAGTGTAGCTGAAAGAAATGTGAAAGAAATCTAGCTCCAACTGTTGAAGAAATCG 2203  
 Qy 6557 GAACCGTTTCCGTAAATCAATTTGCAAGAAAGCAAGAGTAGATGTGAAGAAAGCAAGA 6616  
 Db 2204 TAGCTCCAACTGTTGAAGAAATGTGTAGCTCCAAAGTGTGTAAGAAAGTGTGTCCAAAGTG 2263  
 Qy 6617 TTTTGTGTAAGAAAGTGTAGATTTACAGCAAAACGTTATAGTCTGTGCAATTTGAAATG 6676  
 Db 2264 TTGAAGAAAGTGTGAAGAAAGTGTGAAAGAAAGTGTAGCTGAAAGTGTGAAAGAAAGTG 2323  
 Qy 6677 CCGAGTCCGAGTGGCTGCAAAAGAGCTGGAATTTGAGACAGACAGTGGCAGTTACCAAG 6736  
 Db 2324 TAGCTGAAATGTGTAAGAAAGTGTAGCTGAAAGAAATGTGTAAGAAAGTGTAGCTCCAACTG 2383  
 Qy 6737 ATGAATTCAAACAGAGACAGAGATGTGAAGAAATTTCTAAAT---ATGACTCGAAACAGT 6793  
 Db 2384 TTGAAGAAATTTGTAGCTCCAAAGTGTGAAAGAAAGTGTAGCTCCAAAGTGTGAAAGAAAGTG 2443  
 Qy 6794 TAGATGTAATAGCAAGAAATGAGATTAATCAGGTAAGTGGTTCCAGCCGAGAGCTG 6853  
 Db 2444 TTGCTGAAAGCGTTGCAAAATTTATCAGACAACTTTTAAAGTATTTATTAAGTGTGTA 2503  
 Qy 6854 GAATCTTGACCGAGATATCGAGTGTGTTCTGTCAATAATTTGAAATTAAGTATG 6913  
 Db 2504 TCGAAACTAGGAATTAAGAGACAGTATTTAAATAGATTAAGAGAGTAAGAAAGAAATG 2563  
 Qy 6914 AAACAGATATCATCATAGTATCTTTACACTCTTCTAGATGTAATGTAAGAGCTCTTA 6973  
 Db 2564 TAGTACCAACAATACTAAGAAAGCTGTGAAGAAAGTGTAGCTGAAAGTGTATCTATTTTA 2623  
 Qy 6974 ATTAATTTTCGAATTCCTTGACAGCCGAGTGAAGAGCCGAGGCTCTTGACAGCATTTACCG 7033

Db 2624 GTAAATATTAGAGAGATACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2683  
 Qy 7034 GAGTGTGTTCTGTTAACTAATTAATAGTCTGTGATAGTGTGAGATTGACATAAATCTGTG 7093  
 Db 2684 TAGAAGAACTCCAGAAATATGTAATTAAGTGTGCGCTTTTGAAGAAATACCAAGTGAAGAG 2743  
 Qy 7094 ATTTGACTTCCGTACGAGA---AAAAGTAAATGTACCGCAAAAGAGAAATTAATTA 7150  
 Db 2744 AAAAGAAAGATATAGATGTAATTTGAAGAAAGTAAAGAAAGAGTGTGCTACCACTTTAA 2803  
 Qy 7151 AGCAAAACGACGCAATTCAGAGATTCGAGAGAGCAAGTGGAGCCAAATGCTTTGTAA 7210  
 Db 2804 TAGAACTGTGAAACAGGCAAGAGAAAGACGCAAAATCAATTAACGAAATATTGAAA 2863  
 Qy 7211 ATTAATTTTGAACAGCTGTAGAAATGAAAGAAATTTCTGAAGAAAGAAACAGAGTTT 7270  
 Db 2864 ATTTAGAAAGAAATGCAATAGAAAGTAATGAAGATGTTGCAAGAAATTTAGAGAAATTTAA 2923  
 Qy 7271 TAAAACTTTAGAGAAATTAACAAAGAACAGATTAAGAAAGTAAATGATGCTACGAAA 7330  
 Db 2924 ACGAAACTGTATTTAATCTGTATTTAGATTAAGTAGAGAGAAACAGTAGAAATTTAGCGAG 2983  
 Qy 7331 AATCTTCAATCAGCAGATATTTCTACAGAAAGTACTCTGTAAACCGATAGAGAG 7390  
 Db 2984 AAAGTTTGAAGAAACATGAATATGATTAAGCAATTTTGTAGAAATTTATTAATTTGTA 3043  
 Qy 7391 ATACTCAGAGAGAAAGAAATTAAGCATTTGTGAAGACTCTGTATTAATTTGAAAGAAATG 7450  
 Db 3044 AAGAAATCAAGAAATTTATTAACAGATAG-----TTTGAAGATTAAGAAACAGTA 3097  
 Qy 7451 TAGATATTACAAACAGAGACAAAGATTAATCACTTCTAGCTGTGTGGAACTGAG 7510  
 Db 3098 TAGTATTCAAATCAGAAAGAAAGTGTGATTTGAATGAAAGATGTGTGATTTGATTTAG 3157  
 Qy 7511 GTCTTGTCCGCAATCAGGAACAGTGGCAGTTACAAATTAATTAAGAAATTTCCGAGTTA 7570  
 Db 3158 ATAAATTAAGAAATATGAAAGAAAGTGTATTAATTAATTAATTAATTAATTTTCAAGTACTG 3217  
 Qy 7571 CTGTGAAATTTCTTTTGTGAAGACAGCTGAAAGAAATGTAATGTAGATCGATAT----- 7625  
 Db 3218 AAGGTGTTCAAGAAATCTGTAACATGTGAAACAAATGTATATGTGATGTGATG 3277  
 Qy 7626 ---TACGAAATGTGCTTTAACAGCATATCAAGTCTGTAGAGATTTGGAAATAG 7681  
 Db 3278 TTCCGTATGAAAGATCAATTTTGAAGAAATTTAAAGAGGACAGAGGTTGAAAGAAA 3337  
 Qy 7682 GAGCTGCTATGCAAGATTTAAATTTCTAATGGAAGTCAAAATATCAATTTTAAAAATTTCTA 7741  
 Db 3338 TGTTTTAAATTTGAAAGATGTAATTTAAAGTGAAGTGAATTTACTGTAGAGAAA 3397  
 Qy 7742 AGCTATTGAAAGAAATATGATGTTATTTGTAAGAAATCGGAATTTGAGAGCGGAAG 7801  
 Db 3398 TTAAGATGAACCGGTTCAAAAAGAGGTGAAAGAAAGAACTGTATTTATGAAAGAAA 3457  
 Qy 7802 CAAAAGATTAACCGTAGAGCCGAGTGTGCGGAGCCATTAATCTCAAAAGCAAGAAATG 7861  
 Db 3458 TGAAGAAATATTTAGATGATTAAGAGAAAGAAAGAAAGATTTAACAGACAAAGATGA 3517  
 Qy 7862 AAATGAATTCAGAGTTGAAATTTGAGAAAGATTTTCAATGAAAGAAATAGAGTACTA 7921  
 Db 3518 TAGATCAGTGAAGAAATCCATAGAAATATCTTCAGATTTCAAGAAAGAACTGAAATCTTA 3577  
 Qy 7922 GCCCTTCAAGGAATTTGAAAGAAATCAATGTCAAGTGTGAAAGAAAGAAACAGAGTGA 7981  
 Db 3578 TTAAGATTAAGAAAGAAAGATTTTCACTAGTGTGTTGAAAGAAAGTTCAAGACATGATATGG 3637  
 Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGACAGTGAAGAGGACAG 8027  
 Db 3638 ATGAAAGTGTGAGAAATTTTGAATTTGAAGAAATATGAAAGAGAGATTAATGAAAGATG 3697  
 Qy 8028 AATTAATTTCCAGAGCAAGAGATGCCGAAAGCTCTTATTTGAAAGTATAGTACAAAATCCGG 8087  
 Db 3698 CTGTGAAATTAAGACATTAATGCAAACTTAATGAAAGAACTCAAGAGTTAAATGAAG 3757

Qy	8088	AAGAAGATTTTTCATGCAGATATGGAATGGAAGCAACATTAATAATGAAAGTAC	81.47
Db	3758	TGAAAGCAGATTTTAAATTAAGATATGGAAAAATTTAAAGATTTAGAAAAAGC--ATTATC	3815
Qy	8148	AGCAGTTTCTAAAGCAGTTAACAGGTTCTGTATTTGGGAGAGTTGGAGTCCAAAGCAGA	8207
Db	3816	AGAAGATTCTTAAGAAATTAATAGATGCAAAAAGATATACATTGAAAAAGTTATTGAAGA	3875
Qy	8208	AGTACTGCTGCAGGTAAACTATGTTAGAAAGTTGAGAGAGAAATTTGTTCAACAACA	8267
Db	3876	GGAACTGATTTAACGACGACGTTGGATGAAGTTGTGAATTTAAAGATGTCGAAGAAGA	3935
Qy	8268	TCGATTGAATGCATTTTCTAAAGTAAAGGTTTGGATGGAAGATTAAGTAATCTGTAATC	8327
Db	3936	CAMAGATGAAAAAGTATCTGATTTTAAAGATCTTGAAGAGATATATTTAAAG-----	3988
Qy	8328	TTCTGTAGTATCAGAAATGGAAGGAAATGCGCGAGCAGAGTGAATCTTCTACAGC	8387
Db	3989	--AAGTAAAGAAATCABAAGACTTGAAGTGAATTTTACAGATTTATBAAGATTTAAA	4046
Qy	8388	ACAAAGTATATCTGAATCCGATGTTGCTTTACGAAGCAGATATGAAAAATATATGATTA	8447
Db	4047	AACATTTGAACAGATATTTTAAAGAGAAAAAAGAAATGAAAAAGATCATTTTGAATA	4106
Qy	8448	CACAAAAAATATATTTTCAGAACTCAATGCTCTTCTTTAATGATACAAAGAAATGAGC	8507
Db	4107	ATTGGAAGAAAGAACTGAAAGAAATTAAGAACTTGAAGCAGATATATTTAAAGAAATGAC	4166
Qy	8508	GAATATAGAACTTTAGCGGTAGCCGGTGCATGCACAGGAACAAACAAAGCATTTAC	8567
Db	4167	TTCAATTGAAAGTTGAAGAAAAAATTAAGAAAGATACAGAAATTTAAAGAAAGAGT	4226
Qy	8568	GAGATCAAAACAAGTTAACTTCTACAACTGTAAATGAGAGAAACGATCTCAACTCGTGC	8627
Db	4227	AGAACATATTAATAGTGTGATGCGCATATTAAGAGTTTGGAAAGATGATTTTGAAGA	4286
Qy	8628	AAAAAGCTTTGGCTAAAAATGAAAAATTATGAAATGTAAAGAAACCTGAGGAGCCTTAGT	8687
Db	4287	AGTAGATGA-TTTTAAAGAAAGTATATATGACATGTTTAAAGGAGATATGGA--ATTAGG	4343
Qy	8688	CGAAGCCGAAACAGCAGCCGTTGAAAAATTATACAAAGATATCTACAGAGCATTTGGTTC	8747
Db	4344	GGATATGATATAGGAAGATTGTAGAAAGATTAACAACAAACTTGGAGAAAGATGTAATC	4403
Qy	8748	AGGAATTTGGGAAATTGAGATTAATTGAAACGATTCAGAGATTAATAGATTGTAG	8807
Db	4404	CTTAAGAGATGTTTATATCTAGTGCATTAGCGATGATGAATAACAATGAAAAACAGAA	4463
Qy	8808	AGTCAACGGAACCGAACCAAGAGAGCTTTGTGCGAAAGAGATGTAATTTCTGTGAAAA	8867
Db	4464	AAAAGCTCAAAAGACTAGTGTGGAAGAGTATTTATTAAGAAAGAGTTAAAGAAAGAAC	4523
Qy	8868	TACAAATTTCAAGGGGAAACAAAATCATTCATTGAAGATTAAGCCAGAAATGTTGAAACCG	8927
Db	4524	AAAAGAAAAAATTAACAAAAAAGAAAGTAAGGTTGATTTTAAGATTAAGAACCAAAAAG	4583
Qy	8928	AAAGTAAATATGATGCTGTTGAATGAACCTGATATATCTACAGAAAAAATGGTGG	8987
Db	4584	TGAATATATAGAAAGTTGAATGAAAGATTAAGAAAGATTAAGAAAGATGTAGAAAGAAAT	4643
Qy	8988	CTATGATGGAATTTGGTATTTGGAAAAATGTTGATGTAAATATGATTTAAGAAAAATGTGA	9047
Db	4644	AGAAAGAGTATAGAAAGATTAAGATTGAAAGATTAAGATGAAGATTAATAGTGAAGATAT	4703
Qy	9048	AGCCAAAAATCGAAGACATGCTATTGTAGAACTACGAAAAACAAGATATCAAGCATT	9107
Db	4704	AGGTGAAGCAAAAGATGAAGTATATGATTTAATAGTCCAAAAAGAAACCATTTGAATA	4763
Qy	9108	TACAAAGCAAAAGTAAATATTTCTTGGAAAAAGAGACGCTGACGCTGACGCTCAATATC	9167
Db	4764	GGTTTAACGAAAAAAGAAAAATTTGAAAAAAGTTAAGAAAGGTGTTAAGGTGCTTAA	4823

[illegible]



Db 2612 TTGCTGAAAAAGTTGCAACAAATTTATGACAACTCTTTAAGTAAATTTATTAAGTGGTA 2671  
 Qy 6854 GAATTCCTGACCGGAGATCTGGAGTGGTTCTGTCAATATATTTGCAATTAAGGTG 6913  
 Db 2672 TCGAAACTGAGAAATTAAGGACAGATATTTAAATGAGATAGAGAAAGTAAAGAAATG 2731  
 Qy 6914 AAAAGATATGATCATATGATCTTTACACTCTTTACTGATGTAATGTAAGTAAAGCTCTTA 6973  
 Db 2732 TATGTCACCAATATAGAAAACGTAGAAAACCTACAGCTGAAAGGTGTAACCTATTTTA 2791  
 Qy 6974 ATAAAAATTTGCAATTCCTTGACAGCCGCTGAGAGCCGACAGTCTTGACAGAGTTAACCG 7033  
 Db 2792 GTTACCTATTAAGAGGATACAAAGAAATACATTAATGATGTAATCTAATAGGAAAAAT 2851  
 Qy 7034 GAGTGTCTTCTGTTAACTATTAATATGTTCTGTGATGCTGAGTTCAATTAATCTG 7093  
 Db 2852 TGAAGAACTCCACGAAAATGTAATTAAGTCGCTTAGAAAATACCAAGAGTGAAGAG 2911  
 Qy 7094 ATTTGACTCCGTAACAGA---AAAAGTAAATGTAACGGCAAAAGAGAAAAAATATTA 7150  
 Db 2912 AAAAGAAAGATATGATGTAATTAATTAAGAGTAAAGAGAGGTGCTACCACTTTAA 2971  
 Qy 7151 ACCAAACAGACAAATGACAGATCGAGAGCAGCAATCGAGCCAAATGCTCTTGTTAA 7210  
 Db 2972 TGAATACTGTGAACAGCAGAAAGAAAGAGCCAAATCAATTAACGAAATATTTGAAA 3031  
 Qy 7211 ATTAATTTGGAACAGCTGTAGAAAGATAGAAAAATTTCTGAGAAAAGAACAGAAATT 7270  
 Db 3032 ATTTAGAAAATGACGTAGAAAGATAGAAAATTTGACAGAAATTTTAGAGAAATTTAA 3091  
 Qy 7271 TAAAACTTTAGCGAAGTTAACAAAGAACAGATAAAAAGTAATGATGCTACGAAA 7330  
 Db 3092 ACGAAACTGTATTTAATCTGATTAATTAAGTAAAGTAAAGAACAGTAAATTAAGCGAG 3151  
 Qy 7331 AAATCTTAACATCAGCAGGATTTCTACAGAGATACCTCTGTAAGAAAGCGATAGAGAG 7390  
 Db 3152 AAAAGTTAGAAAACATGAAATGATTAAGCATTTTGTAGTAAATTTGATTAATGTA 3211  
 Qy 7391 ATACTCAGGAGAGAGAAATTAAGCCATTGTGAAAGCTTCTGATTAATTTGAAAAAATG 7450  
 Db 3212 AAGGAATACAGAAATTTATTAACAGTATG-----TTTCGAAATATAGAAACAGTA 3265  
 Qy 7451 TAGATATTACACAGAGACAGAAATATATCTCTTACTGCTGTGTTGGAACTGCGAG 7510  
 Db 3266 TAGTAATCCATCAGAGAAAGAGTGTGATGTAATGTAATGATGTTGATTCATTTAG 3325  
 Qy 7511 GTCTGCTCCGATCAGGAAAGTGGSCGTTACAAATATTAAGAAATTCGCGAGTTA 7570  
 Db 3326 ATTAATATAGAAATATGAAAGAGGTTTATTAATTAATTAAGAAATATTTCAAGTACTG 3385  
 Qy 7571 CTGTTGAAAATCTTTTGTGAAAGCAGCTGAAAAGTAAATGTTAGATCGATAT----- 7625  
 Db 3386 AAGGTGTTCAAGAACTGTATCTGAATCATGTAGAACAAATGTATATGCGATGTTGATG 3445  
 Qy 7626 ----TACAGAAATGTTGCTTTAAACAGCATATCAAGGCTCTGTAGAGCAATTGGAAATG 7681  
 Db 3446 TTCTCTCTATGAAAGATCAATTTTATAGGATATTAATAGAGCAGAGGGTTTAAAGAAA 3505  
 Qy 7682 GAGCTGCTATGAGATTAATTTCTAATGGAAGATCAATATCAGATTAATTAATTTCTA 7741  
 Db 3506 TGTTTTATTTATTTGGAAGATGTATTTAAAGTGAAAGTGTATTAATCTGTAGAGAAA 3565  
 Qy 7742 AGCTATTAAGAAAAATATGATGTTATTTGTAAGATTAATCGAAATGAGACCGAGAG 7801  
 Db 3566 TTAAGATGAACCGGTTCAAAAAGAGGTAGAAAAAGAAACGTGTAATATTAAGAGAAA 3625  
 Qy 7802 CAAAAAGATTAACCGTGTAGAGCGGTAGCTCCGAGCATATCTCAAAAAGCAAAAGATG 7861  
 Db 3626 TGAAGAAATATTTGATGATGTTAGAGAGAAAAAGAAATTTAAACAGACAAAGTGA 3685  
 Qy 7862 AAATGAATTCAGAGTTGAAATTTGAGAAAGATATTTCAATGAAAGAAATAGAGTAACTA 7921  
 Db 3686 TAATGACAGTAAAGAAATCCATAGAAATATCTTCAAGTTCTAAAGAAAGAACTGAATCTA 3745

Qy 7922 GCCCTTAAAGGAATTTGAAAGAGAAATCAATGTCAAGTGGAAAAAGAAAACAGATGA 7981  
 Db 3746 TTAAGATTAAGAAAAAGATGTTTCACTAGTGTGTAAGAAAGTTCAAGCAATGATATGG 3805  
 Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGCAGTATGACGGGCGAG 8027  
 Db 3806 ATGAAAGTGTGAGAAAGTTTATAGATTAAGAAATATGAAAGAGAGATTAATGAGATG 3865  
 Qy 8028 AATTAATTTCCAGCAAAAGATGCCGAAGCTCTTATTTGAAAGTTAGTACAAATCCGG 8087  
 Db 3866 CTGTTGAATTAATGACATTTACTAGCAAACTTATTAAGAAACCTCAAGGTTAAATGAG 3925  
 Qy 8088 AAGAGTATTTTCAATGCAAGTAATATGTAATATGGAACCAACATTAATGAAGTAAC 8147  
 Db 3926 TGAAGCAGATTTAATTAAGATATGAAATTAAGAAATTTAGAAAAAGC--ATTATC 3983  
 Qy 8148 AGCAGTTTCTAAAGCAGTAACAGTTCTGTATTTGGAGGAGTTGGAGTCAACCAAGCAGA 8207  
 Db 3984 AGAAGATTTAAGAAATTAATGATGCAAAAGATGATACATTAAGAAAAAGTTATGAGAG 4043  
 Qy 8208 AGCTACTGCTGAGTAAATCTATGTAGAAAGTTGAGAGAAATTTGTTCAAGACAA 8267  
 Db 4044 CGAATCATATTAACGACGACGTTGAGTGAAGTTAGAAATTTAAAGATGTGAAAGAGA 4103  
 Qy 8268 TCGATTAATGCAATTTCTAAAGTGAAGGTTTGAATGAAGTAAAGTAACTGTTAAATC 8327  
 Db 4104 CAAGATCGAAAAAGTATCTGATTTAAAGATCTTAAGAAAGTATTTAAAG----- 4156  
 Qy 8328 TTCTGTAGATCAGAAATGAGAGGAATTTCCGAGCAGAGAGTGAATCTTCTACAGC 8387  
 Db 4157 --AAGTAAAGAAATCAAAAGAACTTGAAAGTAAATTTAGAAAGTATTAAGAAATTA 4214  
 Qy 8388 ACAAGTAATCTGAATCCGTAGTTGTTACGAAAGCAAGTTATGAAATTAATGATTA 8447  
 Db 4215 AACTATTGAAACAGATATTTTAGAAGAGAAAAAGAAATAGAAAAAGATCATTTGAAAA 4274  
 Qy 8448 CACAAAAAATATATTTCAAGAGTCAATGCTCTCTTAAATGATACAAAGAAATGAAC 8507  
 Db 4275 ATTCGAAAGAGAGCTGAAGAAATTAAGATCTTAAGCAGATATTTAAAGAAAGTATC 4334  
 Qy 8508 GAATATAGAACTTTAGCGGTGCGGTGTCATCAACAGGAACCAACAAAGCAATTTAC 8567  
 Db 4335 TTCAATTAAGAGTTGAAGAAAGAAAAATTAAGAAAGTACACGAATTTAAAGAAAGGT 4394  
 Qy 8568 GAGTCAACAAAGTTAACTTTACAACTGTAAATGAGAGAAAGTATCTCAACTGTCG 8627  
 Db 4395 AGAACATTAATTAAGTGTGATGCCGATTAAGAGTTTGAAGAAAGATGATTTAGAGAG 4454  
 Qy 8628 AAAAGCTTGCTAAATAATGAAATTAATGAAATGTAAGAAAGTGGAGAGCTTACT 8687  
 Db 4455 AGTAGATGA--TTTAAAGAGAGTATATAGACATGTAAAGGAGATATGAG--ATTAG 4511  
 Qy 8688 CGAGCGGAAACAGCAGCCGTTGAAATTAATCAAGAGTACTACAGAGCAATTTGGTTGC 8747  
 Db 4512 GGAATATGATTAAGAAAGTTTGAAGATTAACACAAACCTTGGAGAAAGGTGAATC 4571  
 Qy 8748 AGGAATTTGGGAAATTTGAGATTAATTAAGAAAGATTTGCAAGAGATTAATGATTTAG 8807  
 Db 4572 CTTTAAAGATGTTTTATCTAGTCATTTAGCATGTAGTGAAGAAACAAATGAAACAGAAA 4631  
 Qy 8808 AGTCAACGAGACGAAACCAAGAGAGCTTGTGGAAGAAATGATTTCTGTGAAAAA 8867  
 Db 4632 AAAAGCTCAAAAGCTTAAGTTGAAAGATATTTAAAGAAAGAGTTAAAGAAAGACC 4691  
 Qy 8868 TACAATTTCAAGGGGAAACAAATATCATCTTGAAGATTAAGCCAGATTTGTTGAACCG 8927  
 Db 4692 AAAGAAAAAATTAACAAAAAAGAAATAGTTGATTTAAGATTAAGAAACCAAAAGA 4751  
 Qy 8928 AAGTAAATGATAGATGCTTTGAATGAATCTGATGATTAACAGAAAGAAAGTGTGG 8987  
 Db 4752 TGAATATAGTAAGAGTTGAATGAAAGTGAAGATATTAAGAAAGAGATGTGAAGAGATAT 4811

Qy 8988 CTATGCTGGAATGCTGATTGGAATGATGTAATTAATGCTAATTAAGAAAAATGTA 9047  
Db 4812 ACAAAGAGATATAGAAAGATTAAGTTGATGAATGATGAAGATATATGATGAAGATAT 4871  
Qy 9048 AGCCAAATTCGGAAGACATGCTATTTGTGAAACTACTGAAAAACAAGATATCAAGCAT 9107  
Db 4872 AGGTGAAGCAAAAGATGAAGTTATGATTTAATAGTCCAAAAAGAAACGATTTGAAA 4931  
Qy 9108 TACAAGACAAAGTAATATTTCTTGAAAAAGAGACGCTGCAGCTGCAGTCAATATC 9167  
Db 4932 GGTTAAGCGAAAAAGAAAAATTTGAAAAAAAGTTGAAGAGCGTTAGTGTCTTAA 4991  
Qy 9168 GAATGTACACATTTCCATGAGATGGAATTAATAATTTGGCAAGACATATCATCTTC 9227  
Db 4992 AAAACACGTAGACGAAGTATGATAATATGTTCAAAAAATTTGATTAAGAAAGTTGATTAAGA 5051  
Qy 9228 TCAATTTAATACCAAAAAATTCAAAAATTAATTTACTTTAGCATCA 9274  
Db 5052 AGTATCTTAAGCTTTTGAATCAAAAAATGATGTACTAATGTTTAA 5098

RESULT 21  
AX428966  
LOCUS AX428966 5529 bp DNA linear PAT 21-JUN-2002  
DEFINITION Sequence 1 from Patent EP1201250.  
ACCESSION AX428966  
VERSION AX428966.1 GI:21540347  
KEYWORDS  
SOURCE .  
ORGANISM unidentified  
unidentified  
unclassified sequences.

REFERENCE  
AUTHORS Cohen, J. and Drulhe, P.  
TITLE Immunogenic compositions comprising liver stage malarial antigens  
JOURNAL Patent: EP 1201250-A 1 02-MAY-2002; (BE) ; INSTITUT PASTEUR (FR)  
FEATURES  
source 1.  
1..5529  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="KI Pasteur Strain"

ORIGIN  
Query Match 1.2%; Score 116.2; DB 6; Length 5529;  
Best Local Similarity 41.7%; Pred. No. 4.8e-06;  
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

Db 1301 ATGTGAAAGAAAGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1360  
Qy 5543 AGGCAAAATGTTAAGTTCTACTGAAAAATAGAGTCAGAGGTCAACAAATCGACAGTGC 5602  
Db 1361 GTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1420  
Qy 5603 TTTCGAGCAAGTGAACAACTGCAGTAGAGCTGAGTAGACATTAAATTAATTAATTAAC 5662  
Db 1421 ATGTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAAA 1480  
Qy 5663 AAAATCTCTGACATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5722  
Db 1481 TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1540  
Qy 5723 GC---AAATCTCATTCATTAATAAACAATTTGAAATGAGAGTGTGAGTGTGAGTGTGAG 5779  
Db 1541 ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1600  
Qy 5780 GAGCTGAGTGAACAGGTTCTGTAGCAGTGAATTAAGATTGTAATTAATTAATTAATTAAT 5839  
Db 1601 GTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1660  
Qy 5840 TAAATCATGCAAAAAATCATCTCCGAAGGAAATGTCGAGTTATTACAGAGTCTGATCCGG 5899  
Db 1661 ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1720  
Qy 5900 TAAATGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5959  
Db 1721 TGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAA 1780  
Qy 5960 CCAGTGAATGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6019  
Db 1781 GTGTGTGTAAGAAAGT 1840  
Qy 6020 CTAAAGAAAGAAACAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6079  
Db 1841 GTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1900  
Qy 6080 AAGTATTCAAAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6139  
Db 1901 ATGTGAAGAAAGTGTAGCT-----GAAAATGTTGAAGAAATGTTAGCTCCAACTGTTG 1954  
Qy 6140 AAAAAGATTTGTTTACCAATAGTTCAAGTCTCACTACTTAAATCTTATTTGGCAATG 6199  
Db 1955 AAGAAATGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAAAGTGTGTA---GAAAGTG 2011  
Qy 6200 CCGTGTGTTCAAGACCAAGCCGAGTGTGCAAGAACTGTTAATATCAACAAGTTATGAG 6259  
Db 2012 TGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAAATG 2071  
Qy 6280 AAACAGAGCTCTGTGTAAGAAATCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6319  
Db 2072 TTGAAGAAAGTGTAGCTAATAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATG 2131  
Qy 6320 GAGATTACCAAGAAATTCAGAGTAGAGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6379  
Db 2132 TAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAAAGTG 2191  
Qy 6436 TAGAGCTTCTTCTGATACCAATATTAATAAG---AAATACCAAGCAAGAGTTGAA 6436  
Db 2192 TTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTG 2251  
Qy 6437 AAATCAATGTCGTGATGAAGTTTCCGAGAGAAAGCTGAAATTAACACAGATTCTAAGC 6496  
Db 2252 TAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATG 2311  
Qy 6497 AAGGAATTTCTCTTTTGAAGTCCGAGTCCGAGCAGCGGGGTAGAGCCGAGTGTGAG 6556  
Db 2312 TTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATG 2371  
Qy 6557 GAACGTTTCCGTAAATCAATTTTGCAGAAAGACGAAAGTGAAGTGTGAAGAAAGCAAGA 6616

Db 2372 TAGCTCAACTGTGAGAAATTTGTAAGTCCAAAGTGTGTGTAAGAAAGTGTGCTCCAAAGT 2431  
Qy 6617 TTTTGTAAAAAAGCTGAGATTACAGCAAAAGTTATAGTTCGTGTCAATTGGAATG 6676  
Db 2432 TTGAAGAAAGTGTAGAAAGAAATGTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGT 2491  
Qy 6677 CCGAGTCGGAGTGGCTGCAAAAAGAGCTGGAATTGGAGCAGCAGTGGCAATTCACCAAG 6736  
Db 2492 TAGCTGAAAAATGTGAAGAAAGTGTAGTGAAGAAATGTGAAGAAAGTGTAGCTCCAACTG 2551  
Qy 6737 ATGAATCAAAACAGAGAGCAAGAGTGAAGAAATTTCTAAATTT---ATGACTCGAAACAGT 6793  
Db 2552 TTGAAGAAATTTGAGTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGT 2611  
Qy 6794 TAGATGTAAATAGCAAAAATGAGATTAATCAGGATCTGGAATCGGTCAAGCCGAGCTG 6853  
Db 2612 TTGCTGAAAAAGTGTGCAAAATTTATCAGACATCTTTTAAGTAAATTTATTTAGGGTGA 2671  
Qy 6854 GAATTCCTGCAAGCCGAGATATCTGAGTGTGTTCTGTCAATAATTTGCAAAATTAAGTGA 6913  
Db 2672 TCGAAACTGAGGAATTAAGACAGATATTTAAATGAGATGAAGAAAGTAAAGAAATG 2731  
Qy 6914 AAACAGATATCGATCATAGTACTTTACACTCTTCTACTGATGTAAATGTAAAGCTCTTA 6973  
Db 2732 TAGTACCACTAATCTAGAAAGAGTGAAGAAAGTGAAGTGAAGTGAATCTTTTA 2791  
Qy 6974 ATAAATTTTGAATTCCTTGAACAGCCGAGTGAAGAGCCGAGTCTTGCAGCAGTTACCG 7033  
Db 2792 GTAGCATATTAAGAGGATACAAAGAAATATCTATCTAATGATCTATTAAGAGAAAT 2851  
Qy 7034 GAGTGTCTTCTGTAACTATTAATAGTTCGTGATAGCTGAGTTCACATTAATCTG 7093  
Db 2852 TGAAGAACTCCAGCAAAATGTATTAAGTCCGCTTTAGAAAAATACCAGAAAGTGAAG 2911  
Qy 7094 ATTTGACTCCGTAAGAGA---AAAGTAAATGTAAACGGCAAAAGAGAAAAATATTA 7150  
Db 2912 AAAAGAAAGATTAATAGATGTAAATTTGAAGAAAGTGAAGAGTGTGCTACACTTTTA 2971  
Qy 7151 ASCAAACAGCAGCAAAATGAGAAATCGAGAGCAGCAATCGAGCCAAATGCTTGTGA 7210  
Db 2972 TGAAGAACTGTGAACAGCGAAGAAAGAGCGCAATCAATTAACGAAATTTTGAA 3031  
Qy 7211 ATTAATTTGGAACGCTGTGAAGATGAAGAAATTTCTGAAGAAAGAGCAAGATTT 7270  
Db 3032 ATTTAGAAAGAAATGAGATGAAGAAATGTGCGAGAAATTTAGAGAAATTTAA 3091  
Qy 7271 TAAAACTTTAGAGAGATTAACAAGAACAAATGAATTAAGTGAATGCTACGAAA 7330  
Db 3092 ACGAAAGTGTATTTAATCTGTATTAAGATTAAGTGAAGAAAGTGAAGATTTAGCGGAG 3151  
Qy 7331 AAATCTTAACATCAGCAGGATTTCTACAGAGATCTTCTGTAAGCGGATGAGAG 7390  
Db 3152 AAAGTTTGAAGAAACATGAAGATGATGAAGATTTTGTGAAATATTTGATTAATGTAA 3211  
Qy 7391 ATACTAGGAGAAAGAAATTAAGCCATTGTGAAGCTTGTGATTTATTGAAAAAATG 7450  
Db 3212 AAGGAATACAAAGAAATTTATTAACAGGTATG-----TTTCGAAGATTAAGAAACCAATG 3265  
Qy 7451 TAGATTTTACAGAGAGAAAGAAATTAATCACTTCACTGAGTGTGTTGGAACTGAG 7510  
Db 3266 TAGTAACTCAATCAGAGAAAGAGTGTGATTTGAATGAAGAAATGTGTTAGTTCCATTTAG 3325  
Qy 7511 GTCTGCTTCCGATCAGAAACAGTGGCACTTCAATATTTAAAGAAATTTCCGAGTTA 7570  
Db 3326 ATAAATTAAGAAATTAAGAAAGAGTTTATTAATTAATTAAGAAATATTTCAAGTACTG 3385  
Qy 7571 CTGTGAAATTTCTTTTGAAGAGCAGTGAAGAAAGTAAATGTGAATCGGATAT----- 7625  
Db 3386 AAGGTGTTCAAGAAAGTGTAACTGAAATGTAGAAACAAATGTATATGTGATGTTGATG 3445  
Qy 7626 ----TTCAGGAATATGTTGCTTTAAACAGCATTAAGGTCCTGTAGAGCAGTTGGGAATAG 7681  
Db 3446 TTCTGTATGAAGATCAATTTTGAAGATTTAAATGAAGCAGAGGGTTGAAGAA 3505

Qy 7682 GAGCTGCTATGCAATTTAAATTTCTAATGGAAGATCAAAATATCAGATTTAAAAATTTCTA 7741  
Db 3506 TGTTTTAAATTTGAAAGATGTATTTAAAGGAAAGTGAATTTACTGTGAAGAA 3565  
Qy 7742 AGCTATTAAGAAAAAATATTTGATGTTATTTGAAGAAAGATTAATTCGAATTTGAGAGCGAAG 7801  
Db 3566 TTAAGATGAAGCCGTTCAAAAAGAGGTGAAGAAAGAAAGTGTAGTATTTATGAAGAA 3625  
Qy 7802 CAAAAGATTAACCTGAGAGCGTACTCCGAGCAATTAATCTCAAAAGCAAGAAATG 7861  
Db 3626 TCGAAGAAATATTTGATGATGTATTAAGAGAAAGAAAGAAATTTAACAAGATGA 3685  
Qy 7862 AAATGAATTCGAGGTTGAATTTGAGAGAGATTTTCAATGAAGAAATAGATTAAT 7921  
Db 3686 TAGATGCAATGAAGAAATTCATGAATATTTCTGATTTCTAAAGAAAGAACTGAATCTA 3745  
Qy 7922 GCCCTTTAAAGAAATTTGAAAGAAATCAATGTCAAGAGTGAAGAAAGCAAGTGA 7981  
Db 3746 TTAAGATTAAGAAAGAAAGATGTTTCACTAGTGTGGAAGAAAGTTCAGACATGATATG 3805  
Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGCATGACGGGCGAG 8027  
Db 3806 ATGAAGTGTGAGAAAGTTTAAGATTAAGAAATATGGAAGAGGATTAATGAAGATG 3865  
Qy 8028 AATTAATTCGAGCAAGAAAGATGCCGAAAGCTCTTAATTTGAAGTATGAACAAATCCGG 8087  
Db 3866 CTGTTGAATTAATATACATTAATCAACCTTAATGAAGAAATCTCAAGGTTTAAATGAAG 3925  
Qy 8088 AAGAGTATTTTCTATGCAATGATGTAATGTAAGCAACATTAATGAAGTAAAC 8147  
Db 3926 TGAAGCAGATTTAATTAAGAAATGTAAGAAATTTAAGAAATTTAAGAAAGC--ATTATC 3983  
Qy 8148 AGCAGTTTCTAAAGAGATTAACAGTTCGTATTTGAGAGAGTTGAGTCAACAGCAGA 8207  
Db 3984 AGAAGTTCTAAAGAAATTAATGATGCAAAAGATATACATTAAGAAAGTATTTGAAAG 4043  
Qy 8208 AGCTACTGTCGAGGTAATTAATGATGTAAGTGAAGAGAAATTTGTCAGAACAA 8267  
Db 4044 GGAACATGATTAACGACGATGTAAGTGAAGTGTGAAGTGAAGTGTGAAGAA 4103  
Qy 8268 TCGATTAATGCAATTTCTTAAGTGAAGTGTGTAAGTGAAGTAACTGTAAATC 8327  
Db 4104 CAAGATCGAAAAAGTATCTGATTTAAGAAATCTTGAAGAAATATTTAAAG----- 4156  
Qy 8328 TTCTGTATGATCAGAAATGAGAGAAATTCGAGAGCAAGATGAATTAATCTTAACAG 8387  
Db 4157 --AAGTAAAGAAATCAAGAACTTGAAGTGAATTTTGAAGATTAATGAAGATTA 4214  
Qy 8388 ACAAGTATTAATGAAATCCGTAGCTGTTTGAAGAAAGCAAGATTAAGAAATTAATGAATTA 8447  
Db 4215 AACTATTTGAAGCAAGATTTTGAAGAGAAAGAAAGAAATGAAGAAAGATTTTGAAGAA 4274  
Qy 8448 CCAAAAAATATATTTTGAAGTCAATGCTCTTGTCTTAATGATACAAAGATGAAG 8507  
Db 4275 ATTCGAAGAAAGCTGAAGAAATTAAGATCTTGAAGCAATTAATTTAAGAAAGTATC 4334  
Qy 4335 TTCATTTAGAAAGTTGAAGAAAGAAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAG 4394  
Db 8568 GAGATCAACAGTAACTTCTACAACTGTAATGAGAGAAAGTATCTCACTTGTGTC 8627  
Qy 4395 AGAACATATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4454  
Db 8628 AAAAGCTTGGCTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8687  
Qy 4455 AGTGAATGA--TTTAAAGAAAGTATTTAGACATGTTAAAGGAGATATGA--ATTAG 4511  
Db 8688 CCGAGCAAGAAACAGCAGCCGTTGAAGAAATTAATCAAGAGTATCAAGAGCATTTGTTGTC 8747  
Qy 4512 GATATGATTAAGAAAGTGTGAAGATGTAAACAAAGCTTGAAGAAAGAGTGAATC 4571

Qy	8748	AGGAAATTGGGAAAATTGGAGTAAATTGGAAACGATTGCCAGAGATTAATACGATTGTAG	8807
Dd	4572	CTTAAAGATGTTTATCTAGCGCATTAGGCGATGGTGAACCAATGAAAACAAAGAA	4631
Qy	8808	AGTCACCGAGACGGAACCAAGAGGCTCTGTGCGAAAGATGTAATTTCTGTGAAAA	8867
Dd	4632	AAAAGCTCAAGACCTTAGTTGGAAGAGTATTATTAAAGAAAGGTTAAAGAAAGACC	4691
Qy	8868	TACATATTCAGGGGAAACAAATATCATCTATTGAGATTAAGCCAGAAATTGTTGAAACGG	8927
Dd	4692	AAAGAAAAAATTACAAAAAAGAAAGTAAGTTGATTTTAAGATTAAGAACCAAAAGA	4751
Qy	8928	AAGTGTATGTAGATGCTTGAATGAAGCTTGATGTAGATCTACAGAAAAAGGTGGG	8987
Dd	4752	TGAATATATAGAAAGTTGAATGAAAGATGAAAGATATAGAAAGATGTAGAAAGATAT	4811
Qy	8988	CTATGTGGAATTGTGATTGGAAATGTTGATGTAAATATATGTGATTAAGAAAAATGTAGA	9047
Dd	4812	AGAAAGAAATATAGAAAGAGATTAAGTTGAAGATATAGATAGATATAGATGAAGATAT	4871
Qy	9048	AGCCAAATTCGGAAGACATGCTATTGTGAAACTACGTGAAACCAAGATATCAAGCATT	9107
Dd	4872	AAGTGAACACAAGATGAAAGTTATATTAATAGTCCAAAAGAGAAACGATGGAAAA	4931
Qy	9108	TACAAAGCAAAAGTAAATATTTCTTGGAAAAGAGACGCTCAGCTGCAAGCTGCATATC	9167
Dd	4932	GGTTAAACGAAAAAAGAAAAAATTGAAAAAAGTTGAAAGAGGTGTATGTGCTTTAA	4991
Qy	9168	GAAATGACAAATTCATAGAGATGATATTAATAAATTGGCAAAAGCATGTGATCTTC	9227
Dd	4992	AAACAAGTAGACGAAGTATGAAATATGTTCAAAAAATGATTAAGAAAGTTGATTAAGA	5051
Qy	9228	TCAATTAATTAACCAAAAAATTCAAAAAATATATTAATCTTAGCATCAA	9274
Dd	5052	AGTATCTAAAGCTTAGATCAAAAAAAGATGTTACTATCTTTTAA	5098

RESULT 22				
LOCUS	AX536317	5529 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0238176.			
ACCESSION	AX536317			
VERSION	AX536317.1	GI:25262747		
KEYWORDS				
SOURCE				
ORGANISM	Plasmodium falciparum (malaria parasite P. falciparum)			
REFERENCE	Plasmodium falciparum			
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
TITLE	1			
JOURNAL	Cohen,J. and Druilhe,P.			
FEATURES	Immunogenic compositions comprising liver stage malarial antigens			
source	Patent: WO 0238176-A 1 16-MAY-2002;			
	GlaxoSmithKline Biologicals S.A. (BR) ; INSTITUT PASTEUR (FR)			
	Location/Qualifiers			
	1..5529			
	/organism="Plasmodium falciparum"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:5833"			
	/note="X1 Parasite Strain"			
ORIGIN				

	Query Match	1.2%	Score 116.2	DB 6	Length 5529
	Best Local Similarity	41.7%	Pred. No. 4.8e-06		
	Matches 1747	Conservative 0	Mismatches 2373	Indels 67	Gaps 14
OY	5126	GTTTCGATGAAGACGAGCCTTAGTTGATGATTCGAGTTTGGAAGAGTAAGTCTTTCA	5185		
Db	941	GTTTGAAGAAAGTATAGCTTCAAGTTTGATGAAGATAGTATGATTCAAGTTTAAAGAA	1000		
OY	5186	ATGTGATGACAAAAGATCAAAAACAAATAAATACAA---TTGCGGGAATGCAATGAGG	5242		
Db	1001	ATGTAGCTCCAACTTTGAGAGAAATCGTAGCTCAAGTTGTGTAAAGATGTGCTCCA	1060		
OY	5243	GAAGACGGCTGAGTTTGAGCAACAGTTGCTCATACAAATATTGAAAAACATCAGTTA	5302		

Db	1061	GTGTTGAAGAAAGTGTAAGAAAATGTGAAGAAATGTAGCTGAAAATGTTGAAGAAA	1120
Oy	5303	TAGCTATTGTTAAAAAACAATTAATTAACAACGCGCAATGATCAAGATAGAGAAAAATATCA	5362
Db	1121	GTTAGCTGAAAATGTTGAAGAAAGTGTAGCTAAAAATGTTGAAGAAAGTGTAGCTAAA	1180
Oy	5363	ATGTAGCTGCAAAAAGATTATATCTATGACCAATCTATAGCAGTGGAGTTGGAAGACAA	5422
Db	1181	ATGTTGAAGAAATGTAGCTCCAACGTGTGGAAGAAATGTAGCTCCAACGTGTGAAGAAA	1240
Oy	5423	AAGAGCTCTGTGCAAGAGACTTCTGCAAGTACTACCTTGAAATAGACAGTTTCTTCTC	5482
Db	1241	TTGTAGCTCCAAAGTGTGTAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAA	1300
Oy	5483	ATGTTGATCCAAACGATATTGACCAAAAGATTAGAGAAAGAAATTAATGAAAATGAAGAAA	5542
Db	1301	ATGTTGAAGAAAGTGTAGCTGAAAATTTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA	1360
Oy	5543	AAGCAAAATGTTAATGTTCTAGCTGAAAATACGAGTCAAGTGTGTCACAAATGCGACAGTGC	5602
Db	1361	GTTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTAAA	1420
Oy	5603	TTTCCGAGCAAGTGTGACCAAGCTGCACTAGAGCTGGAAGTGTGACGATTATTAATTAACAC	5662
Db	1421	ATGTTGAAGAAATGTGTAGCTCCCACTGTGTAAGAAATGTGTAGCTCCCAACTGTTGAAGAAA	1480
Oy	5663	AAAATACTCTGACATATTAATAAAATAGTACTCAAAATGTACGAAAATGCTTTGGTAAAAA	5722
Db	1481	TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAA	1540
Oy	5723	GC---AAATCATTCATCTATTAATAAACAATTGGAATTGAGCTGAGATTGGAAGCTGGAG	5779
Db	1541	ATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA	1600
Oy	5780	GAGCTGGAAGTACAGGTTCTGTAGCAGTGAATTAAGATTGAATTAATACATAGACAGAT	5839
Db	1601	GTTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTAAA	1660
Oy	5840	TAAATCATGCAAAAATCACTGCGAAGGGAATATGCGAGTTATTAACAAGTCTGATGCGG	5899
Db	1661	ATGTTGAAGAAAGTGTAGCTGAAAATTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA	1720
Oy	5900	TAAATGCTAATTATGACAGAAACAGTGTCTGAGATGCGCCGTGACAGCAATAGAGACTCAA	5959
Db	1721	TGCTAGCTCCAACGTGTTGAAGAAATGTGTAGCTCCAACGTGTTGAAGAAATGTGTAGCTCAA	1780
Oy	5960	CCAGTGTGAATGAAAATTACAGAGTCTACAAAGCATAATGTAAAGATTCTACAGGTATG	6019
Db	1781	GTTGTTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAA	1840
Oy	6020	CTAAGAAAGAAACAGATGATTATTAATTAATCTCAAGGCAAGTGAATAAGTGTATGATA	6079
Db	1841	GTTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTAAA	1900
Oy	6080	AAGATTCCAAAATCTTAATTAATTAACAAGCTATCACAAAAAGAAAAATTAAGTATATA	6139
Db	1901	ATGTTGAAGAAAGTGTAGCT-----GAAAATGTTGAAGAAATGTGTAGCTCCAACGTGTG	1954
Oy	6140	AAAAAGATTTGTTACCAATAGTTCAGCTACTCACTTAATTTAAATCTTTAATTTGCAAAATG	6199
Db	1955	AAAGAAATCGTAGCTCCAACGTGTGAAGAAATGTGTAGCTCCAAAGTGTGTA---GAAAGTG	2011
Oy	6200	CCGCTGGTTCAGGACCAAGCCGAGTGTGCAGAGAACTGTTAATATCAACAAGTTTATGAG	6259
Db	2012	TGCTGTCAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATG	2071
Oy	6260	AAACAGAAGCTCTGTAGAAAATCTTAATTAATTAATGCAAAAACATTATCTGTAAATACAG	6319
Db	2072	TTGAAGAAAGTGTAGCTGAAAATTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATCG	2131
Oy	6320	GAGATTACAGCAATTCATCGAGTGTAGAGTTCTGTGTGTTGTTGTGAAAATGTAGAG	6379

Db 2132 TAGCTCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGAGCTCCAACTG 2191  
 Qy 6380 TAGGAGCTTCTTCTGATACCAATATTATPAAAAAG---AAATPCCAAGCAAGATTGGAA 6436  
 Db 2192 TTGTGAAGAAAGTGGCTCCAACTGTTGAAGAAAGTGAAGAAATGTTGGAAGAAATG 2251  
 Qy 6437 AAACACAAATGCTGATGAAGGTTTCGGAAGAAAGTGAAGAAATTCAGACAGATCTTAAGC 6496  
 Db 2252 TAGCTGAAGAAATGTTGAAGAAAGTGAAGTGAAGAAATGTTGAAGAAAGTGAAGAAATG 2311  
 Qy 6497 AAGGAATTTCTCTTTTGAAGTCCGAGTCCGACAGCCGGGTAGAGCCGAGTGGCAG 6556  
 Db 2312 TTGAAGAAAGTGTACTGAAAATGTTGAAGAAATGTTGAGCTCCAACTGTTGAAGAAATG 2371  
 Qy 6557 GAACGGTTTCGTAATCAATTTGCAAGAGACGGAAGTATGTGGAAGACAAAGA 6616  
 Db 2372 TAGCTCAACTGTTGAAGAAATTTGAGCTCCAACTGTTGAAGAAAGTGGCTCCAACTG 2431  
 Qy 6617 TTTTGTAAAAAACTGAGATTACAGCAAAAGTTATAGTCTGTGCAATTTGGAATG 6676  
 Db 2432 TTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGAAGTGAAGAAATGTTGAAGAAAGT 2491  
 Qy 6677 CCGAGTCGAGAGTGGCTGCAAAAAGAGCTGAATTTGAGACAGCAGTGGCAATTACCAAG 6736  
 Db 2492 TAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGAAGTGAAGT 2551  
 Qy 6737 ATGAATCAAAACGAGAGACAGAGTGAAGAAATTTCTPAAAT--ATGACTCGAAACAGT 6793  
 Db 2552 TTGAAGAAATTTAGCTCCAACTGTTGAAGAAAGTGAAGTGAAGTCCAACTGTTGAAGAAAGT 2611  
 Qy 6794 TAGATGTAATGACAAAAATGATPAAATCAGTACTGGAAATCGCTTACGCCGAGCTG 6853  
 Db 2612 TTGCTGAAGAAACCTTGCACAAATTTATCGACAACTTTTAAAGTAATTTATTAAGTGGTA 2671  
 Qy 6854 GAATCTTGCAGCCGAGATCTGAGTGGTCTTCTGCAATATATTGCAATTAAGTGA 6913  
 Db 2672 TCGAAACTGAGAAATTAAGAGACAGTATTTAAATGAGTGAAGAAAGTGAAGAAATG 2731  
 Qy 6914 AAACAGATATGATCATAGTACTTTTCACTCTTCTATGATGTAATTTGTAAGCTCTTA 6973  
 Db 2732 TAGTCACCAATATCTAGAAAACGTGAAGAAACCTACAGCTGAAGTGTAACTACTTTTA 2791  
 Qy 6974 ATPAAATTTGCAATCTTGAAGACGCCGGTGAAGAGCCCGAGTCTTGAAGACAGTTACCG 7033  
 Db 2792 GTAACTATTTAGAGAGATCAAGAAATTAATACTAATGATCTATTAAGAGAAAT 2851  
 Qy 7034 GAGTGGTTCTGTAACTATPAAATGTTCTGTAGTACGAGTTCACAAATPACTGTG 7093  
 Db 2852 TAGAAGAACTCCACGAAATGTTATTAAGTCCGCTTTAAGAAATACCCAAAGTGAAGAG 2911  
 Qy 7094 ATTGACTTCCGTACGAGA---AAAAGTAATGTAAACGCAAAAGAGAAATATATTA 7150  
 Db 2912 AAAAGAAAGATTAATGATGTTATTAAGAAAGTAAAGAAAGTGTGCTACACTTTAA 2971  
 Qy 7151 AGCAAAACAGCAAAATGCAAGATCGAGAGACAGCAATCGAGCCAAATGCTTGTGTA 7210  
 Db 2972 TAGAAACTGTGAAACAGCAGAAAGAAAGCGCAATATCAATTAAGGAAATTTTGAA 3031  
 Qy 7211 ATTAATTTGGAACAGTGTGAAGATGAAGAAATTTCTGAAGAAAGAAAGAAAGAGATT 7270  
 Db 3032 ATTTGAAGAAATGCAAGTGAAGATGAAGAAATGTTGCAAGAAATTTAGAGAAATTTAA 3091  
 Qy 7271 TAAAACTTTAGACGAAGTTTAAAGAAACAGATPAAAAAGTAAATGATGCTACGAAA 7330  
 Db 3092 AGCAAACTGTATTTAATCTGTATTAATTAAGTGAAGAAACAGTGAAGAAATTAAGCGAG 3151  
 Qy 7331 AAATCTTACATACAGAGGATTTCTACAGAAATCTTCTGTAAAAAGCGATGAAGAG 7390  
 Db 3152 AAAGTTTGAAGAAATGAAGTGAATGATTAAGCAATTTTGTGAAGAAATTTGATTAATGTA 3211  
 Qy 7391 ATACTCAGGAGAAAGAAATTAAGCAATGTTGAAGCTTCTGATATATTAAGAAAAATG 7450  
 Db 3212 AAGGAATACAGAAATTTATTAACAGTATG-----TTTGAAGTATGAAGAACAGTA 3265

Qy 7451 TAGATATTACACAGAGACAGAAATATACCTTCTACTGTGTGTTGGAACTGAG 7510  
 Db 3266 TAGTAAATCAATCGAAGAAAGTGTGATTTGAATGAAGAAATGTGTAGTTCGATTTAG 3325  
 Qy 7511 GTCTTGTCTCCGATCAGGAACAGTGGCAGTTACAAATTTTAAAGAAATTCGAGTTA 7570  
 Db 3326 ATTAATATGAAGAAATATGAAGAAAGGTTATTAATTAATTAAGAAATTTCAAGTACTG 3385  
 Qy 7571 CTGTGAAATTTCTTTGTGAAGACAGTGAAGAAAGTAAATTTAGATCCGATAT----- 7625  
 Db 3386 AAGGTTCAGAAACGTAACTGAACATGTGAAGAAAGAAATGTATGTGATGTTGATG 3445  
 Qy 7626 ----TACAGAAATGTTGCTTTAACAGCATATCAAGTCTGTGAGACATGGAAATAG 7681  
 Db 3446 TTCTGCTATGAAGATCAATTTTGAAGAAATTTAAATGAGCAGAGAGGTTGAAGAA 3505  
 Qy 7682 GAGTGCCTATGCAAAATTTAAATTTCTAAATGGAAGATCAAAATATCAGTATTTAAAAATCTTA 7741  
 Db 3506 TGTTTTTAATTTGAGAGATGATTTAAAGTGAAGATGATCTAATTTACTGTAAGAA 3565  
 Qy 7742 AGCTATTGAAGAAATTTATGATGTTATTTGAAGATPAAATCGAATTTGAGAGCGAG 7801  
 Db 3566 TTAAGATGAACCGGTCAAAAAGAGTGAAGAAAGAACTGTTGATTTATTTGAAGAA 3625  
 Qy 7802 CAAGAGATTAACGTAAGAGCGTGAAGTCCGAGCCATTAATCTCAAGCAAGAAATG 7861  
 Db 3626 TGAAGAAATTTGTAGATGATTTATTAAGAGAAAGAAAGAAATTTAACAGCAAGATCA 3685  
 Qy 7862 AAATGAATTCAGAGGTTGAATTTGAAGAGATTTTCAATGAAAGAAATAGATPACTA 7921  
 Db 3686 TAGATGCAGTAGAAATATCCATAGAAATATCTTCAGATTTCTPAAAGAGAACTGAATCTTA 3745  
 Qy 7922 GCCCTTGAAGAAATTTGAAGAGAAATCATGTCAAAGTGAAGAAAGCAAGATGA 7981  
 Db 3746 TTAAGATTAAGAAAGAAAGATGTTTCACTAGTGTGAAGAAATTCAGCAATGATATG 3805  
 Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAAGAGCAGTGAAGGCGAG 8027  
 Db 3806 ATGAAGATGTTGAGAAAGTTTGAATTTGAATTTGAAGAAATTTGAAGAGATTAATGAAGATG 3865  
 Qy 8028 AATTAATTTCCGAGCAAAAGATCCGAGAGCTTTATTTGAAGATTAAGTAAATCCG 8087  
 Db 3866 CTGTTGAATTAATGACATTAATACAAACTATTTGAAGAACTCAAGGTTAAATGAAG 3925  
 Qy 8088 AAGAGTATTTTCAATGAGATATGTAATTTGAAGCAACATPAAATGAAGTAAAC 8147  
 Db 3926 TAGAAGCAGATTTAATTAAGATATGAAAGAAATTAAGAAATTAAGAAAGC--ATTATC 3983  
 Qy 8148 AGCAGTTTCAAGCAGTAAACAGTCTGTATTTGGAGAGAGTTGAGTCAACCAAGCGAG 8207  
 Db 3984 AGAAGATTTCAAGAAATTAATGATGCAAAAGATATACATTAAGAAAGTTTGAAG 4043  
 Qy 8208 AGCTACTGTCAGAGTAAATCTATGATGAAGTGTGATGAAGTAAAGTAACTGCTAAATC 8267  
 Db 4044 GGAACATGATTAAGACAGACGTTGATGAAGTGTGAAGTAAAGATGTCGAAGAG 4103  
 Qy 8268 TGTGTGATGATCAGAAATGAGAGAGATTTCCGAGACAGATGAATTAATCTTACAGC 8270  
 Db 4104 CAAGATCGAAAGATCTGATTTAAAGATCTTGAAGAGATATTTTAAAG----- 4156  
 Qy 8328 TTCTGTGATGATCAGAAATGAGAGAGATTTCCGAGACAGATGAATTAATCTTACAGC 8387  
 Db 4157 --AAGTAAAGAAATCAAGAACTTGAAGTGAATTTTGAAGATTTAAGAAATTTAA 4214  
 Qy 8388 ACAAGTATTAATCTGAATCCGATGTTGTTAGAAAGCAAGATTAAGAAATTAATGATTA 8447  
 Db 4215 AACTATTGAAGACAGATTTTGAAGAGAAAGAAAGAAATTAAGAAATTAATGATTA 4274  
 Qy 8448 CACAAAGAAATATATTTCAAGATCAATGCTGCTTTTAAATGATACAAAGATGAAGC 8507  
 Db 4275 ATTGGAAGAGAGAGTGAAGAAATTAAGATCTTGAAGCAGATATTTTAAAGAGATATC 4334

QY 8508 GAATATAGATCTTTAGCCGTAGCCGGTGTGATGCACAGAACAACAAACATTTAC 8567  
DB 4335 TTCTATTAGAGTTGAAGAGAAAAAATTTAGAGAGTACACGAATTTAAAGAGAGGT 4394  
QY 8568 GAGATCAACAGATTAACCTTCAACACTGTAATGAGAGAACTGATCTCACTTCGTGC 8627  
DB 4395 AGAACATATATAGTGTGATGCGCATTAAGAGTTTGGAGAGAGATGATTTGAGAGA 4454  
QY 8628 AAAACCTTGGCTAAATAATGAAATTTATGGAATGTAAAGAGACTGAGAGACCTTACT 8687  
DB 4455 AGTACATGA- TTTAAAGAGAGATATATTAGACATGTTAAAGAGAGATATGA- -ATTAG 4511  
QY 8688 CGAGCCGGAACAGCAGCCCTGTTGAAATATACAAAGAGTACTACAGAGCATTTGGTGC 8747  
DB 4512 GATATGATTAAGAGAGTTTGAAGATGTACCAACAAACTTGAAGAGAGGTGAATC 4571  
QY 8748 AGGAATTTGGGAAATTTGGAGATTAATTGAAACGATTCGACAGATTAATACATTTAG 8807  
DB 4572 CTTAAAGAGTGTATTATCTAGTGCATTAGGCATGATGAAGACAAATGAAAAACAAGAA 4631  
QY 8808 AGTCAACGAGACGAGAACCAAGAGTCTTGCAGAAAGATGGTATTTCTGTGAAAAA 8867  
DB 4632 AAAACCTCAAAACCTTAGTGGAGAGATATTATTTAAAGAGAGTTAAAGAGAAC 4691  
QY 8868 TACAAATTCAGGGGAAACAAATTCATCCATTGAAGATAAGCCAGATTTGTGAACCG 8927  
DB 4692 AAAGAAAAAATTAACAAAAAAGAGTGTATTAAGATTAAGAGAACCAAAAGA 4751  
QY 8928 AAGGTAAATGTAGATGCTTGAATGAACCTTGAATGATCTACAGAGAAAAAGGTGG 8987  
DB 4752 TGAATATGAGAGTTGAAATGAAGATGAAGATTAAGAGAGAGTGAAGAGATAT 4811  
QY 8988 CTATGTGGAATTTGATTTGGAATGTTGATGTAATATGATTTAAGAAAAATGTAGA 9047  
DB 4812 AGAAGAGATTAAGAGAGAGATTAAGTTGAGATTAAGATGAAGATTAAGAGAGATAT 4871  
QY 9048 AGCCAAATTCGAGAGACATCTATTGTAAGAACTACTGAAAAAGAGATATCAACATT 9107  
DB 4872 AGGTGAACAGAAAGATGAAGTATTAATTAAGTCCAAAAAGAGAACGATGAAAA 4931  
QY 9108 TACAAGAGCAAAAGTAATATTTCTTGAAAAAGAGACCTGACGTGCAGCTGCATATC 9167  
DB 4932 GGTAAAGGAAAAAGAAAAATTAGAAAAAAGTTGAAGAGGTGTTAGTCTTAA 4991  
QY 9168 GAATGTACACATTTCCAAATGAGATGATTTAAATTTGGCAAGCAGTATGACTTTC 9227  
DB 4992 AAAACAGTGAAGCAAGTAATGAATATGTTCAAAAAATGATTAAGAGAGTTGATTAAGA 5051  
QY 9228 TCAATTATTAACCAAAAAATTTCAAAAAATTAATTAATTAATTAATTAATTAATTA 9274  
DB 5052 AGTATCTAAGCTTGAATCAAAAAATGATGTACTAATGTTTAA 5098

RESULT 23  
AS8932 6152 bp DNA linear PAT 06-MAR-1998  
LOCUS AS8932 6152 bp DNA linear PAT 06-MAR-1998  
DEFINITION Sequence 1 from Patent WO9641877.  
ACCESSION AS8932  
VERSION AS8932.1 GI:3714406  
KEYWORDS  
SOURCE  
ORGANISM  
unidentified  
unclassified  
unclassified sequences.  
REFERENCE  
1 Drullhe, P. and Daubersies, P.  
AUTHORS MALARIU, PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
TITLE PATENT: WO 9641877-A 1 27-DEC-1996;  
JOURNAL PASTEUR INSTITUT (FR)  
COMMENT Other publication AU 6309696 970109  
Other publication FR 2735478 961220.  
FEATURES  
Location/Qualifiers  
1..6152  
/organism="unidentified"

ORIGIN  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"

Query Match 1.2%; Score 116.2; DB 6; Length 6152;  
Beet Local Similarity 41.7%; Pred. No. 4,6e-06;  
Matches 147; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

QY 5126 GTTCTGATGAAACGAGACCTTTAGTTAGTGAATTCGAGTTTGAAGAGTAAATCTTCA 5185  
DB 1017 GTGTTGAAGAAAGTATACCTTCAAGTGTGATGAAGAAAGTATAGATTTCAAGTATGAAGAAA 1076  
QY 5186 ATGTAGATGCAAAAGATCAAAACAAATAAATACA--TTGCCGAAATGCAAAATGAG 5242  
DB 1077 ATGTAGCTCAACGTGTGAAGAAATCGTAGCTCCAAAGTGTGTAGAAAGTGTGCTCCAA 1136  
QY 5243 GAAAACGGCTGAGATTGAGACACAGTTGCTCATTAATTTGAAAAACAATCAGTTA 5302  
DB 1137 GTTGTGAAGAAAGGTAGAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAA 1196  
QY 5303 TAGCTATTGTAAAAACAGTAAATTTACACGGCGAATGATCAAGATGAAGAAAAATATCA 5362  
DB 1197 GTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGAAA 1256  
QY 5363 ATGTGACTGCAAAAGATTATATCTATGACCAATCTATAGCAGTCGAGTTGAGAGAGCA 5422  
DB 1257 ATGTGAAGAAATGTGACTCCAACTGTGAAGAAATCGTAGCTCCAACTGTGAAGAAA 1316  
QY 5423 AAGAGCTCTGTGCAAGAGCTTCTGCAGTACTCTTGAATTAAGCAGTTTCTTC 5482  
DB 1317 TTGTAGCTCCAAAGTGTGTAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAA 1376  
QY 5483 ATGTGATCAACATGATATTGACAAAGATTTAGAGAGAGAAATTAATGAAATAAGAAA 5542  
DB 1377 ATGTGAGAAAGAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAA 1436  
QY 5543 AAGCAATGTTAATGTTCTAGCTGAAAATGACAGTCAGAGTGCACAAATGAGCAGTGC 5602  
DB 1437 GTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAA 1496  
QY 5603 TTTCCGAGCAAGTGAACAGCTGCAATGAGAGCTGAGAGTACAGTTAATAAATTACAC 5662  
DB 1497 ATGTGGAAGAAATGTAGCTCCAACTGTTGAAGAAATGTGTAGCTCCAACTGTTGAAGAAA 1556  
QY 5663 AAAATACCTTCGACATTAATAAATAGTACTCCAAATGTGCGAAATGCTTGTGTAATAA 5722  
DB 1557 TTGTAGCTCCAAAGTGTGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAA 1616  
QY 5723 GC--AAATCTCATTCATCTATTAACAAATTGGAATTTGAGCTGAGAGCTGAGAG 5779  
DB 1617 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAA 1676  
QY 5780 GAGCTGAGTACAGAGTTCTGTAGCAGTGAATTAAGATTGTAAATATATCATAGCAGAAAT 5839  
DB 1677 GTGTAGCTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAA 1736  
QY 5840 TAAATCATGCAAAATATCATCGCAAGGGAATGTGCGAGTTATTACAGAGCTGATGCGG 5899  
DB 1737 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAA 1796  
QY 5900 TAAATGCTAATTATGAGAGAAACAGTGTGAGAGTCCCGTGCAGCAATAGAGCTCA 5959  
DB 1797 TGTAGCTCCAAAGTGTGAAGAAATCGTAGCTCCAACTGTGAAGAAATTTAGTCTCCAA 1856  
QY 5960 CCAAGTGAATGAATTAACAGATCTACAAAGCATATGTAAAGATTCTACAGTGAATG 6019  
DB 1857 GTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAA 1916  
QY 6020 CTAAGAGAGAAACAGATGATTAATTAATCTACTCAAGGCAAGTGTGATTAAGTGAAGATA 6079  
DB 1917 GTGTAGCTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAA 1976  
QY 6080 AAGTATTCAAAAATCTTAATTAATTAACGAGAGCTTATCACAAAAAAGAAAAATTAAGTATA 6139

Db 1977 ATGTGGAAGAAAGTACT-----GAAATGTGGAAGAAATGCTACTCCAACTGTG 2030  
Qy 6140 AAAAAGATTTGTACCAATAGTTCAGCTACTCATACTTTAAATCTTTATTTGGCAATG 6199  
Db 2031 AAGAAATCGTAGCTCCAACTGTGGAAGAAATGTAGCTCCAAAGTGTGTA---GAAAGTG 2087  
Qy 6200 CCGCTGTTTCAGGACAAAGCCGAGTGCAGGAACCTGTTAATTCACAAGTTTATGAG 6259  
Db 2088 TGGCTCCAAAGTGTGGAAGAAAGTGAAGAAAGTGTGGAAGAAAGTGTAGCTGAAAATG 2147  
Qy 6280 AAAAGAAAGCTCTGTAGAAAATTCATATTATATGCAAAACATATTCTGTAAATCAG 6319  
Db 2148 TTGAAGAAAGTACTGAAATGTTGAAGAAAGTACTGAAATGTTGAAGAAATCG 2207  
Qy 6330 GAGATTACAGAAATCAATCCGAGTAGAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGAG 6379  
Db 2208 TAGCTCCAACTGTGGAAGAAATGTAAGCTCCAACTGTGGAAGAAATGTTAGCTCCAAAGTG 2267  
Qy 6380 TAGAGCTTTCTGTATACCAATTTATTAAG---AAATACCAAGACAGATTGGAA 6436  
Db 2268 TTGTGAAAGTGTGCTCCAAAGTGTGGAAGAAAGTGTGGAAGAAATGTTGGAAGAAAGTG 2327  
Qy 6437 AAACATCAATCTGATGAAAGTTTCGAGAAAGTGAATTTACAGCAGATTCTAAGC 6496  
Db 2338 TAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAAATG 2387  
Qy 6497 AAGGAATTTCTCTTTTGGAGTCGAGATCGCAGACCCGAGTAGAGCCGAGATGGCAG 6556  
Db 2388 TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATGTAAGCTCCAACTGTGGAAGAAATG 2447  
Qy 6557 GAAACGTTTCCGTAATCAATTTGCAGAAAGCGAAGTAGATGTGAAGAACCAAGA 6616  
Db 2448 TAGCTCCAACTGTGGAAGAAATGTAGCTCCAAAGTGTGGAAGAAAGTGTGCTCCAAAGTG 2507  
Qy 6617 TTTTGGTAAAAAAGCTGATTAACGCAAAACGTTATAGTCTGTGCAATTTGGAATG 6676  
Db 2508 TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTG 2567  
Qy 6677 CCGCAGTCGAGTGCCTGCAAAAGAGAGCTGGAATTTGAGCAGCAGTGCAGTTAACCAAG 6736  
Db 2568 TAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTCCAACTG 2627  
Qy 6737 ATGAATCAAAACGAGAGCAAGAGTGAATAATTTCTAAAT---ATGACTCGAACAAGT 6793  
Db 2628 TTGAAGAAATGTAGCTCCAAAGTGTGGAAGAAAGTGTAGCTCCAAAGTGTGGAAGAAAGTG 2687  
Qy 6794 TAGATCTATACAGAAATGAGATTAATCAGTACTGGAATCGGTTCAAGCCGAGCTG 6853  
Db 2688 TTGCTGAAAGCGTTGCAAAATTTATCAGACAATCTTTTAAAGTAATTTATTAAGTGTGA 2747  
Qy 6854 GAATTTCTGAGCCGAGATATCGAGTGTCTGCAATATATTTGCAAAATTAAGTAG 6913  
Db 2748 TCGAAACTGAGAAATTAAGGACAGTATTTAAATGAGATGAAGAAAGTGAAGAAAGTG 2807  
Qy 6914 AAACAGATATCGATCATAGTACTTTAACACTCTTTAATGATTAATTTGTAAGTCTTTA 6973  
Db 2808 TAGTCACCAATACTAGAAACGTAAGAAACCTACAGCTGAAAGTGAATCTACTTTTA 2867  
Qy 6974 ATTAATTTTGAATTTCTTGAACGCGGTGAGAGACCCGAGTCTTGACGAGTTAACG 7033  
Db 2868 GTTAACATATTAAGAGATTAAGAAATTAATTAATTAATGATCTATTAAGAGAAAT 2927  
Qy 7034 GAGTGTCTTCTGTTAACATATTAATAGTTCGTGATAGCTCGAGTTCAATTAAGTCTG 7093  
Db 2928 TAGAAGAACTCCAGAAATATGATTAAGTCCGCTTTAAGAAATACCCAAAGTGAAGAG 2987  
Qy 7094 ATTGAATTCCTGACAGAG---AAAAGTAATTTAAACGCAAAAGAGAAATTAATTA 7150  
Db 2988 AAAAGAAAGAAATTAATGATGATTAATGAAGAAAGTGAAGAGTGTGCTACCACTTTAA 3047  
Qy 7151 AGAAACAGAGCAATTCAGGAATCGAGAGACAGCAATCGAGCAATGTCTTGTA 7210

Db 3048 TAGAAACTGTGAAACAGCAGCAAGAAAGACCGCAAAATACATTTACGGAATATTGAAA 3107  
Qy 7211 ATAAATTTGAAACAGCTGTAGAAAGATGAAGAAATTTGAGAGAAAGCAAGATTT 7270  
Db 3108 ATTTAGAAAGAAATGCGATGAAAGATGAAGAAATGTTGCAAGATTTTAAAGAAATTA 3167  
Qy 7271 TAAAACTTTAGAGAAATTAACAAGAACAGATTAAGAAATTAAGTAATGATCTACGAAA 7330  
Db 3168 ACGAAATCTGATTTTAATCTGTATTAAGTAAGTAAGAGAAACAGTAAATTTAGCGAG 3227  
Qy 7331 AAATCTTAACAATCAGCAGATATTTCTACAGAAAGATCTCTGTAAACCGATAGAGAG 7390  
Db 3228 AAAGTTTGAAGAAACAAATGAATGATTAAGCAATTTTATGCAATATTTATTAATGTA 3287  
Qy 7391 ATACTCAGAGAAAGAAATTAAGCAATTTGGAAGACTCTGATTAATTTGGAAGAAATG 7450  
Db 3288 AAGGAATCAAGAAATTTAATTAACAGGTAG---TTTCGAAGTATAGAAACAGTA 3341  
Qy 7451 TAGATATTACAACAGAGCAAGAAATATATCACTTCTACTGTGTTGGGAACTGAG 7510  
Db 3342 TAGTAAATCAATCAGAGAAAGATGTTGATTTGAATGAAGAAATGTGTTAGTTCATTTAG 3401  
Qy 7511 GTCTTGCTCCGATCAGAAACAGTGCAGTTACAAATATTAAGAAATTTCCGAGTTA 7570  
Db 3402 ATAAATATGAAGAAATATGAAGAGGTTATTAATTAATTAAGAAATTTTCAAGTACTG 3461  
Qy 7571 CTGTTGAAATTTCTTTGTGAAGACAGCTGAAAGATTAAGTTAGATCCGATAT---- 7625  
Db 3462 AAGGTGTCAAGAACTGTAACTGAACATGTAGAAACAAATGTATATGTGATGTTGATG 3521  
Qy 7626 ----TACGAAATGTTGCTTTAAACAGATATCAAGTCTGTAGAGACATGGAAATG 7681  
Db 3522 TTCTGCTATGAAGATCAATTTTGAAGATTTTGAAGATTTAAAGACGAGAGGTTGAAGAAA 3581  
Qy 7682 GAGCTCCATGACAGAAATTAATTTAAATGGAAGATCAATATCAGTATTAATAATTTCTA 7741  
Db 3582 TGTTTTTAATTTGGAAGATGATTTAAAGGAAGTGAATTAATTAATCTGTAGAGAAA 3641  
Qy 7742 AGCTATTGGAAGAAATATGATGTTATTTGTAAGATTAATTCGAATTTGAAGCGGAG 7801  
Db 3642 TTAAAGATGAACCGGTTCAAAAGAGTGAAGAAAGAAACGTTAGTATTAATTTGAAGAAA 3701  
Qy 7802 CAAGAGATTAACGTAAGAGCGGTAGCTCCGAGCCATTAATCTCAAAAGCAAGAAATG 7861  
Db 3702 TGAAGAAATATTTATGATGTTATTAAGAGAAAGAAAGATTTTAACAGCAAGATGA 3761  
Qy 7862 AAATGAATTCAGAGGTTGAAATTTGAAGAGTATTTCAATGAAGAAATTAAGTAACTTA 7921  
Db 3762 TAGATGAGTGAAGAAATCCATGAATAATCTTCAGATTTCTAAAGAACTGAATCTTA 3821  
Qy 7922 GCCCTTTAAGGAATTTGGAAGAAATCAATGTCMAAGTGAAGAAAGAAACAGATGA 7981  
Db 3822 TTAAAGATTAAGAAAGATGTTTCACTAGTGTGGAAGAAATTTCAAGCAATGATATG 3881  
Qy 7982 CTGCGAATTCAGAGAGCT-----TCTGTAAGAGCAGTGAAGGCGCAG 8027  
Db 8028 AATTAATTTCCAGAGCAAGAAATGATCCGAGATCTTAATTTGAAAGTTAGTACAAATCCG 8087  
Qy 3942 CTGTTGAATTAATGATCATTTAAGCAAACTTAATGAAGAACTCAAGAGTTAAATGAAG 4001  
Db 8088 AAGAGTATTTTCAATGCAATATGTAATTTGAAGCAACATTAATTAATGAAGTAAAC 8147  
Qy 4002 TAGAAGCAATTTAATTAAGATGATGAAGAAATTAAGAAATTTAGAAAAAGC--ATTATC 4059  
Db 8148 AGCAGTTCTAAGCAGTAACAGGTTCTGATTTGGAGGAGTTGAGTCAACCAAGCAG 8207  
Qy 4060 AGAAGATTTTAAGAAATATGATGCAAAAGATGATACATTAAGAAAGTATTAAGAA 4119  
Db 8208 AGTACTGCTGAGAGTAAATTAATGATGTAAGTGAAGAGAAATTTGTTAGAAACAA 8267  
Qy 4120 GGAACATGATTAAGACGACGTTGATGAAGTGTGAATTTAAAGATGTGCAAGAGAA 4179

8268 TCGATTGGAATGCAATTTCTAAAGTAGAAAGTTGGATGAGATTAAGTAATGCTAAATC 8327  
4180 CAAAGATCGAAAAAGATATCTGATTTAAAGATCTTGAGAGAAATATTTAAAG----- 4232  
8328 TTCTGTAGTATCAGAGAAATGAGAGAAATTCGCCGAGCAGAGATGTAATCTTACAGC 8387  
4233 --AAGTAAAGAAATCAAAAGAACTTGAAAGTAAATTTTGAAGATTTATTAAGAAATTA 4290  
8388 ACAAGTAAATCTGAATCCGTAGTTCGTTTACGAAGCAAGTTATGAAAAATATGATTA 8447  
4291 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATGAAAAAGATCATTTGAAAA 4350  
8448 CACAAAAAATATATTTTCAAGATCAATGCTCTCTTAAATGATATACAAAGATGAAGC 8507  
4351 ATTCGAGAGAGAAAGCTGAAGAAATTAAGATCTTGAGAGAGATATTTAAAGAAATATC 4410  
8508 GAATATGAAATCTTTAGCCGTAGCCGTTGTCATGCACAGAGACAAACAAAGCATTTAC 8567  
4411 TTCAATTGAAAGTTGAAGAAAGAAAAAATTTGAAGAGTACAGAAATTTAAAGAGAGT 4470  
8568 GAGATCAACAGATTAACCTTCAACATGTAATGAGAGAAACGTATCTCACTTCGTGC 8627  
4471 AGAACATATATTAAGTGTGATGCGCATATTAAGGTTTGAAGAGATGATTTGAAGA 4530  
8628 AAAAGCTTTGCTAAAAAATGAAATTTATGAAATGTAAGAGAACTGAGAGACCTTAGT 8687  
4531 ACTATAGTGA-TTTAAAGAGAGATATTTAGACATGTTAAAGGAGATATAG--ATTAG 4587  
8688 CGAGCGGAAACAGACCCGTTGAAATATATCAAGAGTACTACAGAGCATTTGTTGC 8747  
4588 GGAATATGATTAAGAAAGTTGAAGATGTAAACAAACTTGAAGAAAGATTTGAATC 4647  
8748 AGGAATTTGGGAAATTTGAGATTAATTTGAAGATGCAAGAGATTAATACATTTGTAG 8807  
4648 CTTAAAGATGTTTATCTAGTCAATGATGAGTGAAGACAAATGAAAAACAAGAA 4707  
8808 ACTCAACGAGACGGAACCAAGAGAGTCTTGCAGAAAGATGATTTCTGAAAAA 8867  
4708 AAAAGCTCAAAACCTTAAGTTGAGAAATATTTAAAGAGAGGTTTAAAGAAAGACC 4767  
8868 TACAAATTCAGGGGAAACAAATTCATCCATGTAAGATTAAGCCAGAAATTTGGAACCG 8927  
4768 AAGGAAAAAATTAACAAAAAAGAAAGTATTAAGATTAAGATTAAGAACCAAAAGA 4827  
8928 AAGTGAATGTAATGCTTTGAATGAATGTAATGTAATCTACAGAGAAAGTGTGC 8987  
4828 TGAATATGTAAGATTTGAATGAATGAATGAATGAAGAGATGTAAGAGAAATAT 4887  
8988 CTATGCTGAATTTGATATGAAATGTTGATTAATAATGATTAAGAAAAATGTAGA 9047  
4888 AGAAGAGATATGAAGAGATTAAGTTGAAATATAGTGAAGATTAAGTGAAGATAT 4947  
9048 AGCCAAATCGAGAGACATGCTATTGTAGAAATCTAGAAAAACAAGATTAACAGATT 9107  
4948 AGGTAAAGCAAGATGAAGTATAGATTTAATAGTCAAAAAAGAAACGATTTGAAAA 5007  
9108 TACAGAGCAAAAGTAAATTTCTTGAAAGAGACGTCGACGTCGACATATC 9167  
5008 GGTAAAGGCAAAAAAGAAAAATTTGAAAAAAAGTTGAAGAGGTTTGAAGTCTTAA 5067  
9168 GAATGTACACATTTCCATGAGATGATATTAATAATTTGGCAAGACATATGATCTTC 9227  
5068 AAAACGATGAAGAGATATATGTTCAAAAAATTTGATTAAGAGATGATTAAGA 5127  
9228 TCAATTAATACCAAAATTTCAAAATATATTAATTTAGTCATCA 9274  
5128 AGTATCTAAAGCTTTAGATCAAAAAATGATGTTACTATATGTTTAA 5174

RESULT 24  
AR131271  
LOCUS AR131271 6152 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 1 from patent US 6191270.  
ACCESSION AR131271  
VERSION AR131271.1 GI:14119596  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 6152)  
AUTHORS Drullhe P. and Dauberies P.  
TITLE Malarial pre-erythrocytic stage polypeptide molecules  
JOURNAL Patent: US 6191270-A 1 20-FEB-2001;  
FEATURES  
source location/Qualifiers  
1..6152  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 1.2%; Score 116.2; DB 6; Length 6152;  
Best Local Similarity 41.7%; Pred. No. 4.6e-06;  
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

5126 GTTCGATGAACCGAAGCTTTAGTATGATTTCTGAGTTTGAAGAGTAACTTCTCA 5185  
1017 GTTTGAAGAAAGTATAGCTTCAAGTGTGATGAAGATTCAGATTTGAAGAA 1076  
5186 ATGTAGATGCAAAAGATCAAAACAAATTAATACAA--TTGCCGAAATGCCAATGAG 5242  
1077 ATGTAGCTCCAACTGTTGAAGAAATCGTACCTCCAGTGTGTGAAGAGTGTGCTCAA 1136  
5243 GAAAAAGCGCTGAGTGTGAGCAACAGTTGCTCATCAAAATTTGAAAAACATCAGTTA 5302  
1137 GTGTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAGTGAAGAA 1196  
5303 TAGCATTTGTAAAAACAGTAAATTAACAGCGGAAATGATCAAGATGAAAAATATCA 5362  
1197 GTGTAGCTGAAAGTGTGAAGAAAGTGTGAAAGTGTGAAGAAAGTGTAGCTGAA 1256  
5363 ATGTAGCTGCAAAAGATTAATTAATTAATGACCAATTAATGACAGTGTGAGAGAGCA 5422  
1257 ATGTGAAGAAATGTAGACTCCAACTGTTGAAGAAATGTAGCTCCAACTGTGAAGAA 1316  
5423 AAGAGCTCTGTGCAAGAGACTTCTGCAGTACTACCTGAATTAAGACAGTTCTTCTC 5482  
1317 TTGTAGCTCCAAAGTGTGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAGAA 1376  
5483 ATGTGATCAACAGTATTTGACAAAGATTTTGAAGAGAAATTAATGAAATTAAGAA 5542  
1377 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTGAAGAA 1436  
5543 AGCCAAATGTTAATGTTTACCTGAAATTAAGAGTCAAGTGTGCAATGCGACAGTGC 5602  
1437 GTGTAGCTGAAGAAATGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAA 1496  
5603 TTTCCGAGCAAGTGAAGACAGTGAAGAGTGAAGAGTGAAGAGTGAATTAATTAACAC 5662  
1497 ATGTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGTAGCTCCAACTGTTGAAGAA 1556  
5663 AAAATATCTTGCACATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5722  
1557 TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA 1616  
5723 GC---AAATTCATTCATCTATTAACAATTTGAATTTGAAGTGTGAAGTGTGAG 5779  
1617 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAA 1676  
5780 GAGCTGAGTGAAGAGTCTGTAGAGTGAATTAAGATTTGAATTAATTAATTAATTAATTA 5839  
1677 GTGTAGCTGAAGAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAA 1736  
5840 TAAATCATGCAAAATTAATCACTGCGAAGGAAATGTTCGAGATTAATTAATTAATTAATTA 5899  
1737 ATGTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAAGTGTAGCTGAAGAAATGTTGAAGAAA 1796

QY	5900	TAATTTGCTAATTATGACGAAACAGTGTCTGGAGTGGCCCGTCGACCAATAGAGCTCGAA	5959
Db	1797	TCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAA	1856
QY	5960	CCAGTGTGAATGAATAATTAACAGATCTCAAAAGCATATGTAAAGATTTCTACGTGATG	6019
Db	1857	GTTGTTGTAAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAA	1916
QY	6020	CTAAAGAAAGAAACAGATGATTAATTAATTAACACTCAAGGGCAAGTATGATTAAGTGTAGATA	6079
Db	1917	GTTAGAGCTGAAATATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGTGA	1976
QY	6080	AAGTATTCAAAATCTTAATTAATTAACGAGAATTATCACAAAAAAGAAAAATAATATA	6139
Db	1977	ATGTTGAAGAAAGTGTAGCT-----GAAAAATGTTGAAGAAATGTTAGCTCCAACTGTTG	2030
QY	6140	AAAAAGATTTGTTATCCAAATAGTTCAAGTACTCACTACTTTAAATCTTTATGTGCAATG	6199
Db	2031	AAGAAATCGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAAAGTGTGTA--GAAAGTG	2087
QY	6200	CCGCTGGTTGAGACAAGCCGAGCGGAGCAAGCTGTTAATATCAACAAGGTTATAGAG	6259
Db	2088	TGGCTCCAGGTGTTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTAGCTAAATG	2147
QY	6260	AAACGAAGCTCTGTGAGAAATTTCTAATTAATTAATGCAAAAATTATTTCTGTAAATCAG	6319
Db	2148	TTGAAGAAAGTGTAGCTGAAATATGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCG	2207
QY	6320	GAGATTACAGAAATTCAATCGAGTAGTAGTTCTGTTGGTGTGGTGAATGTAGAG	6379
Db	2208	TAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAAAGTG	2267
QY	6380	TAGAGCTCTTCTGATATCAATTAATTAATAAG--AAATACCAAGCAAGAGTTGAA	6436
Db	2268	TTGTGAAGAAAGTGTAGCTCCAACTGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTG	2327
QY	6437	AAACTACAAATGTCTGATGAGAGTTTCGAGAGAAGAGCTGAATTAACAGCAGATTTCTAAGC	6496
Db	2338	TAGCTGAAATATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAATG	2387
QY	6497	AAGGAATTTCTCTTTTGAAGTCGAGTCGACAGCCGGGTGAGACCGGAATGGCAG	6556
Db	2388	TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATGTAGCTCCAACTGTTGAAGAAATCG	2447
QY	6557	GAACTGTTCCGTAATCAATTTGACAGAAAGACGAAAGTAGATGTGAAAGAACCAAGA	6616
Db	2448	TAGCTCCAACTGTGAAAGAAATGTAGCTCCAAAGTGTGAAAGAGTGTGCTCCAAAGTG	2507
QY	6617	TTTTGCTAAAAAAGCTGAGATTACAGCAAAACGTTATAGTTCGTGTCCAATTTGGAATG	6676
Db	2508	TTGAAGAAAGTGTAGAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTG	2567
QY	6677	CCGAGTCGAGTGGCTCAAAAGAGCTGGAATTTGACAGACAGTGGCACTTACAAAG	6736
Db	2568	TAGCTGAAATATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTCCAACTG	2627
QY	6737	ATGAATCAAAACGAGAGCAAGTGAAGAAATTTCTAAATTT---ATGACTGGAACAAGT	6793
Db	2628	TTGAAGAAATGTGTAGCTCCAACTGTTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTG	2687
QY	6794	TAGATGTATAAGCAAAATGAGATTAATCAAGTACTGGAATCGGTTCAAGCCGAGCTG	6853
Db	2688	TTGCTGAAAAAGTGTGCAACAAATTTATCAGACAACTTTTAAGTAAATTTATAGTGGTA	2747
QY	6854	GAAATCTTGACGCCGAGTATCTGAGTGTCTGTCAATATATTTGCCAATAAAGTAG	6913
Db	2748	TCGAAACTGAGAAATTAAGACAGATATTTAAATGAGATAGAGAAATTAAGAAAGTAATG	2807
QY	6914	AAACAGATTCGATCATATGACTTTTAACTCTTTACTAGATGTAATGTTAAAGCTCTTA	6973
Db	2808	TAGTACCAACAATATCTAGAAAAAGTGAAGAAACTACAGCTGAAGTGTAACTCTTTTA	2867
QY	6974	ATPAAATTTTCAATTTCTTGAACAGCCGTTGAGAGACCGCAGTCTTGACAGAGTTTACG	7033

Db	2868	GTAACTATTAGAGAGATCAAGAAATCTATTACTAATGATCTATTGAGAAAAAT	2927
Oy	7034	GAGGTGTTCTGTTAACAATAAATAGTCTGTGATGCTGAGTTCACAATACTGTG	7093
Db	2928	TAGAAAGAACTCAAGAAATGATTAAAGTCCGCTTTAGAAAAATCCAAAGTGAAGG	2987
Oy	7094	ATTTGACTCCGTCGAGA---AAAGTAAATGTACGGCAAAAGAGAAAAATTTA	7150
Db	2988	AAAAAGAAAGATATGTATGTATTTGMAAGTAAAAAGAGAGTGTACCACTTTAA	3047
Oy	7151	AGCAACAGACGCAAAATGCAGGATCGAGAGAGCAATCCGAGCAATGTCCTGTAA	7210
Db	3048	TAGAACTGTGGAAACAGACGAGAAAAAGCCAAATCAATTACGAAATTTTGA	3107
Oy	7211	ATAATTTTGGAAACGCTGTAGAAATGAAAAATTTCTGAAGAAAAAGAACAGATTT	7270
Db	3108	ATTTAGAAGAAATGAGTAGTAAGAAATGTTGACAGAAATTTAGAGAAATTA	3167
Oy	7271	TAAAACTTTAGACGAAGTTAACAAAGAACAGATAAAAAAGTAATGATGCTACAAA	7330
Db	3168	ACGAACCTGATTTTAACTGATTTAATTAAGTAGAGAAACGTAGAAATTTGCGAG	3227
Oy	7331	AAATCTTACATCAGCAGTATTTCTACAGAAATCTTCTGTAAAGCGATAGAGAG	7390
Db	3228	AAAGTTTAGAAACATGAATGATTAAGCATTTTGTAGAAATTTTGATTAATGTA	3287
Oy	7391	ATACTCAGGAGAAAGAAATTAAGCCATGTGAGACTTGTATATTGGAaaaaATG	7450
Db	3288	AAGAAATCAAGAAATTTATTACAGTATG-----TTTCGAATATAGAAACAGTA	3341
Oy	7451	TAGATATTACACGAGAGCAAGATTAATCATCTTCTACGTGCTTTGGAACTCGAG	7510
Db	3342	TAGTAATCCATCGAAGAAAGGTTAATTTGAATGAAAGTGTAGTTGATTCGATTTT	3401
Oy	7511	GTCCTGCTCCGCAATCAGAACAGTGCAGTTACAAATATTAAAGAAATCCGAGTTA	7570
Db	3402	ATAATATAGAAATATGAAAGAGTTTATTAAATATTAGAAATATTTCAAGTACTG	3461
Oy	7571	CTGTTGAAATCTTTGTGAAAGCAGCTGAAAAAGTAATGTTAGATCGATAT-----	7625
Db	3462	AAGGTGTTCAAGAACTGTAACTGAACATGTAGAACAAATGTTATGTGATGTTGATG	3521
Oy	7626	-----TACAGAAATGTTGCTTTAACAGATATCAAGGCTCGTGAAGCATTTGGAAATG	7681
Db	3522	TTCCTGATAGAAAGATCAATTTTTAGAATATTAAATAGAGCAGAGGGTTGAAGAA	3581
Oy	7682	GAGCTGCTATGCAGAATTAATCTAATGAGAAATCAAAATATCAGTATTTAAAAATCTA	7741
Db	3582	TGTTTTTTAATTTGGAGATGTATTAAAGTGAAGATGTAATTACTGTAAGAAA	3641
Oy	7742	AGCTATTAGAAAAAATATGATGTTATTGTAAGATTAATTCGAATTGAGCGGAG	7801
Db	3642	TTAAGATGAAACCGGTTCAAAAAGAGTAGAAAAAGAACTGTATGATTATTGAAAGAA	3701
Oy	7802	CAAAAGATTAACCGTAGAGACCGGTAGTCCGAGACCATATCTCAAAACCAAGATG	7861
Db	3702	TGAAAGAAATATGTATGATGTATTAGGAGAAAGAAAGATTTAACACAGATGA	3761
Oy	7862	AAATGAATTCAGAGTTGAAATTTGAGAGATATTTTCATAGAGAAATAGATTACTA	7921
Db	3762	TAGATGCAGTAGAAATCCATAGAAATATCTTCAGATTTAAAGAGAACTGAATCTA	3821
Oy	7922	GCCCTTTAAAGAAATGGAAGAGAAATCAATGTCAAAAGTGAAGAAAAAGAAACAGATGA	7981
Db	3822	TTAAAGATTAAGAAAGAAATGTTTCACTAGTGTGGAAGAGTTCAACACATGATATGG	3881
Oy	7982	CTGCTGAATTCACAGAGCT-----TCTGTAGAGCAGTAGACGGGCGAG	8027
Db	3882	ATGAAGTGTGAGAAAGTTTTTAGAATTTGAAAAATATGGAAGAGAGATTAAGAAAGATG	3941
Oy	8028	AATATTTCCGAAGAAAGATGCCGAGCTCTTATTGAAAGTATGATCAAAATCCG	8087

Db 3942 CTGTTGAAATTAATGACATTACTACGAACTTATTGAAAGAACTCAAGGTTAAATGAG 4001  
 Qy 8088 AAGAATATTTTTCAGCAGATATGTAATGAGAGCAACATTAATGAAGTAAAC 8147  
 Db 4002 TAGAAGCAGATTTTAAAGATATGGAATAATTTAAAGAAATTAGAAAAAGC-ATTATC 4059  
 Qy 8148 AGCAGTTTCTAAAGCAGTACAGGTTCTGTATTGGAGAGGTTGAGTCACCAAGCCAGA 8207  
 Db 4060 AGAAGATTTCTAAAGAAATTAATGATGCAAAAAGATGATCATTTAGAAAAAGTTATTGAAAGA 4119  
 Qy 8208 AGCTACTGTCGAGGTTAAACTATGGTGAAGTTAGAGAGAAATTTGTCAGAACAA 8267  
 Db 4120 GGAACATGATATTAACGACGACCTGGATGAAGTTGATTAAGATTTAAAGATGTCGAAGAAGA 4179  
 Qy 8268 TCGATTGATGCAATTTCTTAAAGTAGAAGGTTGATGAGATGAATAAGTAACTGCTAAATC 8327  
 Db 4180 CAAGATTCGAAAAAGTATCTGATTTTAAAGATCTTGAAGAAAGTATATTAAAG----- 4232  
 Qy 8328 TTCTGTAGTATCAGAAATGAGAGAGAAATGCCGAGCAGAGTGAATCTTCTACAGC 8387  
 Db 4233 --AAGTAAAGAAATCAAAAGAACTTGAAGTAAATTTTGAAGATTTTAAAGATTTAA 4290  
 Qy 8388 ACAAGTATATCTGAATCCGTAGTTCGTTTACGAAAGCAAGATTATGAAAAATATGATTA 8447  
 Db 4291 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATGAAAAAGATCATTTTGAAGA 4350  
 Qy 8448 CACAAAAAATATATTTTCAGAGTCATGCTCTGCTTAAATGATATACAAAGATGAGAC 8507  
 Db 4351 ATTGCAAGAGAGAGCTGAGAAATTAAGATCTTGAAGCAGATATATTAAAGAAAGTATC 4410  
 Qy 8508 GAATATGAAATCTTTAGCCGTAGCCGCTGTGATGCAAGAGAACAAAGCAATTTAC 8567  
 Db 4411 TTCATTAGAGTGAAGAAAGAAAAAAATTAAGAGATGACGAAATTTAAAGAAAGAGT 4470  
 Qy 8568 GAGATCAACAAAGTTAACTTCAACAGCTTAATGAGAAAGATATCTCACTTCGTCG 8627  
 Db 4471 AGAACATTAATTAAGTGGATGCGCATTAAGAGTTTGAAGAAAGATGATTTGAAGA 4530  
 Qy 8628 AAAAGCTTTGCTTAAATGAATTAATGAAATGTAAGAGAACTGAGAGACCTTACT 8687  
 Db 4531 AGTATATGA- TTTAAAGAGAACTATATTAGACATGTTAAAGAGATATGGA-ATTAG 4587  
 Qy 8688 CGAGCCGAGAACAGACGCCGTTGAAATTAACAAGAGTACTACAGAGACATGCTTGC 8747  
 Db 4588 CGATATGATTAAGAAAGTTTGAAGATGTAACAACAACTTGAAGAAAGATGGAATC 4647  
 Qy 8748 AGGAATTTGGGAAATTTGAGATTAATTTGAAACGATGCAAGAGTAAATACATTTAG 8807  
 Db 4648 CTTAAAGATGTTTTATCTAGTCATTTAGCATTTGATGAGAAACAAATGAAAAACAAGAA 4707  
 Qy 8808 AGTCAACGAGACGGAACCAAGAGAGGCTTGTCCGAAAGATGATTTCTGTGAAAAA 8867  
 Db 4708 AAAAGCTCAAAAGCCTTAAGTTGAGAAAGTATTTAAAGAAAGGTTTAAAGAAAGACC 4767  
 Qy 8868 TACAAATTCAGGGGAAACAAAATCATCTGAGAGATAAGCCAGAAATTTGTGAACCG 8927  
 Db 4768 AAAGAAAAAAATTAACAAAAAAAGTAAGTTGATATTAAGATTAAGAAACCAAAAGA 4827  
 Qy 8928 AAGTAAATGTAAGTCTTGAATGAAATTTGATGTAATGTAATGTAATGTAATGTAATGTA 8987  
 Db 4828 TGAATTAAGTAAAGTGAATTAAGATGAAGATTAAGAAAGATGTAAGAAAGATAT 4887  
 Qy 8988 CTATGTGGAATTTGATTAAGAAATGTTGATGTAATTAATGATTAAGAAATGTTAGA 9047  
 Db 4888 AAGAAAGATTAAGAAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 4947  
 Qy 9048 AGCCAAATTCGAGAGACATGCTATTGTAGAAACTACTGAAAAACAAGATTAACAGAT 9107  
 Db 4948 AGGTGAAGCAAAAGATGAAGTATAGATTTAATAGTCAAAAAAGAAACGATTTGAAAA 5007  
 Qy 9108 TACAAAGCAAAAGTAAATATTTCTTGAAGAAAGACGCTGACGCTGACGCTCAATATC 9167  
 Db 5008 GGTAAAGCAAAAGAAAAATTTAGAAAAAAAGTTGAAGAGGTGTTAGTGTCTTAA 5067

Qy 9168 GAATGTACACATTTCCATGATGATGATATTTAAAAATTTGGCAAGACGATATGATCTTC 9227  
 Db 5068 AAAACAGCTAGACGAAGTAAGTAATATGTTCAAAAAATGATTAAGAAAGTTGATTAAGA 5127  
 Qy 9228 TCATTAATTAACCAAAAAATTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTA 9274  
 Db 5128 AGTATCTTAAGCTTTAGAAATCAAAAAAGATGTTACTAATGTTTAA 5174

RESULT 25  
 AC164627  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

AC164627 192275 bp DNA linear ROD 27-JUL-2005  
 Mus musculus 10 BAC RP23-382A18 (Rosewell Park Cancer Institute  
 (C57BL/6J Female) Mouse BAC library) complete sequence.  
 AC164627  
 AC164627.4 GI:71274204  
 HTG.

Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muroidae; Muridae; Murinae; Mus.

REFERENCE  
 AUTHORS

1 (bases 1 to 192275)  
 Muzny, D., Adams, C., Agdai II, O., Allen, C., Albrooks, S., Archer, P.,  
 Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,  
 Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Burch, C., Burch, P.,  
 Cadore, J., Canada, A., Cardenas, V., Carter, K., Cavazos, J.,  
 Chacko, J., Chahour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,  
 Cheng, M., -T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,  
 Curry, S., Dai, W., Davila, M. L., Davis, C., Davy, Carroll, L., De  
 Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,  
 Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durkin, K.,  
 Dziuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,  
 Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,  
 Fowler, G., Fu, Q., Fun, B., Garcia, A., Garcia, R., Garner, T.,  
 Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D.,  
 Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, M., Heberlein, K.,  
 Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Hawlak, P.,  
 Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,  
 Hines, S., Hitchens, M., Hodgson, A., Hogue, M., Hollins, B.,  
 Howell, L., T., Huylk, S., Hume, J., Imo, K., Jackson, A., Jackson, L.,  
 Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalatus, K., Kelly, S.,  
 Keys, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F.,  
 Leal, S., Lee, K., Lee, S., Legall, F. I., Lemon, S., Lewis, L., Li, B.,  
 Li, Y., Li, Z., Linell, M., Liu, W., Liu, Y., -S., Liu, Y., Liyange, D.,  
 London, P., Lopez, J., Lorenshewa, L., Lorado, R., Luk, T., Mach, R.,  
 Maheshwari, M., Mahoney, C., Malloy, K., Mansouri, D., Martinez, B.,  
 McClelland, H., McPherson, J., Mercado, C., Metzker, M.,  
 Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Munsida, M.,  
 Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Bastan, B.,  
 Nott, A., Nwaelele, O., Obregon, M., Ochli-Okorio, C., Odeh, E.,  
 Okunolu, G., Okunolu, K., Parker, D., Pasternak, S., Patel, B.,  
 Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Phan, T.,  
 Primus, E., Pu, L., -L., Puzo, M., Qin, X., Quinn, A., Quitor, J.,  
 Rabata, D., Rachlin, B., Reigh, R., Ren, Y., Reuter, A., Richards, S.,  
 Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sama, M., Sanders, W.,  
 Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y.,  
 Sisson, I., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R.,  
 Svalter, A., Taylor, T., Thomas, N., Thorm, R., Thornton, R.,  
 Trejos, Z., Umami, K., Varjo, C., Verdusco, D., Villaseana, D., Virk, D.,  
 Volkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J.,  
 Wei, X., Wheeler, D., Williams, G., Williams, R., Morley, K., Wright, R.,  
 Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L.,  
 Zhang, Z., Zhou, D., Weinstock, G. and Gibbs, R.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE

2 (bases 1 to 192275)  
 Morley, K. C.  
 Direct Submission  
 Submitted (25-JUN-2005) Human Genome Sequencing Center, Baylor  
 College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 192275)



Db 11197 AGAGCAGAGCAAGAGCAGCAGCAACAGAGCAGAGCAAG 11256  
Qy 5670 TTCTGCACATATATAAATAAGTACTCAAAATGTACGAATTCGTTTGTATAAAGCAATC 5729  
Db 11257 AAGACGACAGAGAGCAGCAGCAGCAACAGAGCAGCAGCAGCAACAGAGCAGC 11316  
Qy 5730 TCATTCTATATTAATAAATTTGAAATTTGAGCTGAGTTGAGCTGAGAGCTGAGT 5789  
Db 11317 AGAGCAGCAGAGTGTACACAGAGCAGCAGTAAACAGAGCAGCAGAGCAGCTGCTC 11376  
Qy 5790 GACAGCTTGTAGCAGTGAATTAAGATTTTAATATATACAGATGAGCAATTAATCATGTC 5849  
Db 11377 AGCAGAGAGAGCAGAGGAGGAACAGTGTAAACACAGCAGCAGCAACAGAGCAGC 11436  
Qy 5850 AAAAATCACTGCGAAGGAGAAATGTCCGAGTTATTACAGAGTCTGATTCGGTATTTGCTAA 5909  
Db 11437 AACAGAGCAGCAGCAGCAGCAACAGAGCAGCAGCAGTAAACAGAGCAGCAGAGCAGC 11496  
Qy 5910 TTATGCAAGAACAGTGTCTGAGGTGCCCCGTGCAAGCAATAGAGCCTCAACCAAGTGTGA 5969  
Db 11497 AGTAGTAGCAACAGAGCAGCAGCAACAGAGCAGCAGAGCAGCAGCTCCAGCAGCAGC 11556  
Qy 5970 TGAATATTAAGGATCTCAAAAGCATATGTAAAGATTTTACAGTGTATTGCTAAAGAGA 6029  
Db 11557 AGCAACAGCAACAGAGCAGCAGCAACAGAGAACACAGAGCAGCAGAGAGCAACAGT 11616  
Qy 6030 AACAGATGATTAATTAATTAATTAATCAAGGCAAGTAGATTAAGTGTATTAAGTATTCAA 6089  
Db 11617 GGTAAACACAGCAGCAGCAGCAACAGAGCAGCAGCAGAGAGCAGCAGCAACAGAGCAGC 11676  
Qy 6090 AAATCTTAATTAATTAACGAAGACTTATCACAAAAAGAAAAATAGTATTAATAAAGATTT 6149  
Db 11677 AGGTCCAGAGGTCCAGCAGGTCCAGCAGCAGCAGCAGCAGCAGAGCAGCAGCAACAGCA 11734  
Qy 6150 TGTTCACCAATAGTTCAGCTACTCATCTTTAAATTTTATTTGGCAATCCCGCTGTTC 6209  
Db 11735 -GAAGCAACAGAGAGCAGAGGAGGAACAGTGTAAACACAGCAGCAGCAGCAACAGAGC 11793  
Qy 6210 AGGACAAAGCCGAGAGTGGCAGGAACCTGTAAATCAACAAGTTTATGAGAAACAGAAAC 6269  
Db 11794 AGCAGCAGAGCAGCAGAGAGGAACAGTGTAAACACAGCAGCAGCAGCAACAGAGCAGC 11853  
Qy 6270 TCTTGTAGAAAAATCTATATTAATGAATGCAAAACATTAATCTGTAAATCAGAGATTAAC 6329  
Db 11854 AGCAGAGCAGCAGCAACAGAGCAGCAGAGTCCAGCAGGTCCAGCAGCAGCAGCAACAG 11913  
Qy 6330 GAATTCATTCGAGTAGTAGTGTCTGTGTGTGTGTGTGAATGTAGAGTAGAGCTTC 6389  
Db 11914 AGCAGAGCAACAGTAGAAGCAACAGAGGAGCAGAGAGAACAGTGTAAACAGAGCAGC 11973  
Qy 6390 TTCTGATACCAATATTAATAAAGAAATACCAAGACAAAGATTGAAAAATACATGTC 6449  
Db 11974 AGCAGCAACAGCAGCAGCAGAGCAGCAGAGAGAACAGTGTAAACAGCAGCAGC 12033  
Qy 6450 TGAAGAGTTTCGAGAAAGAGCTGAAATTAACAGCAGATTCTAAGCAAGAAATTTCTC 6509  
Db 12034 AGCAACCAAGAGCAGCAGAGAGCAGCAGCAACAGAGCAGCAGAGTCCAGCAGCTCAGC 12093  
Qy 6510 TTTTGAAGTTCGAGTGCAGCAGCCGGGTAGAGC-----CGAGTGGCAGAAACCT 6563  
Db 12094 AGCAGAGCAACAGAGCAGCAGCAACAGAGAACAGAGGAGGAGCAGAGAGAACAT 12153  
Qy 6564 TTCCTGAATCAATTTGAGAAAGACGGAAGTAGATGTGAAGAAAGCAAGTTTGT 6623  
Db 12154 GGTAAACAAGCAGCAGCAGCAACAGAGCAGCAGAGAGCAGCAGAGCAACAGTGT 12213  
Qy 6624 AAAAAAGCTGAGATTACAGCAAAACGTTATAGTTCTGTTGCAATTTGAAATGCGCAGT 6683  
Db 12214 AACCAACAGCAGCAGCAGCAACAGAGCAGCAGAGAGCAGCAACAGAGCAGCAG 12273  
Qy 6684 CGAGTGTGCTGCAAAAGAGCTGGAATTTGAGAGCAGTAGTGCGATTTACAAAGATGAATC 6743  
Db 12274 TCCAGAGGTCCAGCAGCAGCAACAGAGCAGCAGCAACAGAAACCAACAGAGGGA 12333

Qy 6744 AACACGAGACCAAGTGAATAATTTCTAATATATGACTCGAAAACAATTAGATGTAT 6803  
Db 12334 GCAGAG--GAACAGTGTAAACAACAGCAGCAGCAACAGAGCAGCAGAGCAGCAGC 12392  
Qy 6804 AGCAGAAAAATAGATTAATTAAGTACTGGAATCGGTTCAGCCCGAGCTGGAATTTCTTC 6863  
Db 12393 AGGAGGAACAGTGTAAACAACAGCAGCAGCAACAGAGCAGCAGCAGAGCAGCAGC 12452  
Qy 6864 AGCCGAGTATCTGAGAGTGTCTGTCAATATTTGCAATATAAGTAGAAACAGATAT 6923  
Db 12453 AACAGAGCAGCAGGTCCAGAGGTCCAGCAGCAGCAGCAACAGAGCAGCAGCAACAGC 12512  
Qy 6924 CGATCATAGTACTTTACACTCTTCTAGATGTAATGTAAAGCTCTTAATTAATTTTC 6983  
Db 12513 AGAAGCAACAGAGCAGCAGAGGAACAGTGTAAACAACAGCAGCAGCAACAGAGCAGC 12572  
Qy 6984 GAATTCCTTGAACGCGGTGAGAGCCGCAAGTCTTGCAAGCATTAACCGAGTGTTC 7043  
Db 12573 GGCAGCAGAGCAGCAGCAACAGAGCAGCAGCTCCAGCAGCAGCAGCAAC 12632  
Qy 7044 TGTTAACCTAATTAATAGTCTGTGATGCTCGAGTTCAATTAATCTGATTTGACTTC 7103  
Db 12633 AGAGCAGCAGCAACAGCAGAAACAGAGAGCAGAGAGAACAGTGTAAACAACAGC 12692  
Qy 7104 CGTACGAGAAAAAGTAAATGTAAACGCAAAAGAGAAAAATTAAGCAACAGCAGC 7163  
Db 12693 AGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 12752  
Qy 7164 AAATGCAAGAAATCGAGAGCAGCAATGAGGCCATGTCTTGGTAAATTTTGGAAAC 7223  
Db 12753 AGCAGAGAGCAGCAGCATGAGCAGCAGAGCAGAGCAGCAACAAATATCAGCAGAGC 12812  
Qy 7224 AGCTGTAGAATGAAAAAATTTGAGAGAAAAAGGAACAGAGTTTAAAAAATTTAGA 7283  
Db 12813 AATGAACACACAGCAGAGGAAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAACAGC 12872  
Qy 7284 CGAAGTTAAACAAAGAACAGATTAATAAAGTAATGATCTAGAAAAAATCTTACAATC 7343  
Db 12873 AACAGCAACTGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGTAGC 12932  
Qy 7344 AGCAGTATTTCTACAGAAATCTTCTGTAAAGCGAGTAGAGAGTACTCAGAGAGA 7403  
Db 12933 AGCAGCAGCATGAGCAGCAGTGAACAAATATCAGCAACCAACAGAGCAGCAGCAG 12992  
Qy 7404 AGGAATTAAGCAATTTGAGAGCTTGTGATATTAATGGAATAATGATATTACAAC 7463  
Db 12993 AGAAGCAGCAGAGCAGCAGCAGCAGCAGAGCAACAGAGCAGCAGCAGCAGCAAC 13052  
Qy 7464 AGAGACAAAGATATATCACTTCACTGTGTGTGTGTGTGGAACCTGAGGTCTTGCTCCGC 7523  
Db 13053 CAGAGCAGCAAGAAAGCAGAAATCAACAGAGAGCAGCAGAGAGAGTGTAAACAGCAGC 13112  
Qy 7524 ATCAGAAACAGTGCAGTTACAAATATTAAGAAATTCGAGATTCTGTGAAAAATTC 7583  
Db 13113 AGCAGCAGCAGCAACCAAGCAGCAGAGAGCAGCAGCAGCAACAGCAACAGCAGC 13172  
Qy 7584 TTTTGTGAACAGCTGAAAAAGTAAATGTTAGATCGGATTTTACAGGAATTTGTGCTTT 7643  
Db 13173 AGCAGCAGCAGCTGAGCAGCAGAGAGTAGAGCAGCAGCAGCAGCAGAGAACAGAGC-- 13229  
Qy 7644 AACAGCATATCAAGTCTGTAGAGCATTTGGGAATAGAGCTGCTTATGCAATTTAA 7703  
Db 13230 AGCAGAGCAGCAGTCCAGCAGCAGCAGCAACAGAGACACAGCAACAGCAGAAACAA 13289  
Qy 7704 TTTTAATGAAGATCAATATCAGTATTAATAATTTCTAAGCTATTAGAAAAAATTTGA 7763  
Db 13289 AGAAGAGCAGAGAGAACAGTGTAAACAACAGCAGCAGCAGCAACCAAGAGCAGCAG 13349  
Qy 7764 TGTATTTGTAAGATTAATCGGAATTTGAGCCGGAACCAAAAGATTTAACCGTAGAGC 7823  
Db 13350 AGCAGCAGCAAGAGCAGCAGAGAAAGCAGCAGAGACAGCAGAGAACCTTAGAGAGCAGC 13409



```

* 170587 170686: gap of unknown length
* 170687 170685: contig of 6279 bp in length
* 176966 177065: gap of unknown length
* 177066 187042: contig of 9977 bp in length
* 187043 187142: gap of unknown length
* 187143 203395: contig of 16253 bp in length.
Location/Qualifiers
1. 203395
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-16N3"
53146..53245
/estimated_length=unknown
159348..159447
/estimated_length=unknown
162069..162168
/estimated_length=unknown
166875..166974
/estimated_length=unknown
170587..170686
/estimated_length=unknown
176966..177065
/estimated_length=unknown
187043..187142
/estimated_length=unknown

```

## ORIGIN

Query Match 1.2%; Score 113.8; DB 14; Length 203395;  
Best Local Similarity 41.3%; Pred. No. 2.6e-06;  
Matches 1153; Conservative 0; Mismatches 1617; Indels 19; Gaps 5;

```

QY 5610 AGCAAGTGACAAAGCTGACAGTAGAGCTGAGTAGAGCTTAATTAATTAACAAATATAC
DB 137742 AGGAGAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 5670 TTCTGCACATTAATAAATAGTACTCAAAATGTACAAAATGCTTTGGTAAAAAGCAATC
DB 137682 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 5730 TCATTCATCTATTAATAAATGGAATTTGGAAGTGAATTTGGAAGTGAAGTGAAGTGA
DB 137622 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 5790 GACAGGTTCTGTAGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATC
DB 137562 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 5850 AAAAATCACTCGAAGGAAATGTGGAAGTTTACAGAGCTGATGCGTAATTGCTTA
DB 137502 AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 5910 TTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 137442 AGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 5970 TGAATTAACAGATCTACAAAAGCATATGTAAAGATTCTACAGTATGCTAAAGAGA
DB 137382 AGCAACAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6030 AACAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
DB 137322 GGTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6090 AAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
DB 137262 AGGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6150 TGTACCAATAGTTCAGTACTACTCTTTAAATCTTTATTGGCAATGCGCTGGTTC
DB 137204 -GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6210 AGGACAAAGCCGAGTGGAGAGAACTGTAAATATCAACAAAGTTATGAGAGAAACAGAG

```

```

DB 137145 AGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6270 TCTTTAGAAAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
DB 137085 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6330 GAATTCATCGAGAGTAGAGGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
DB 137025 AGCAGCAGCAACAGTAGAAGGCAACAGAGGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6390 TTCTGATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
DB 136965 AGCAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6450 TGATGAAGGTTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136905 AGCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6510 TTTTGAAGTCGAGAGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136845 AGCAGCAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6564 TTCTGATTAATCAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136785 GGTAAACACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6624 AAAAAGAGCTAGATTAACAGCAAAACGTTATAGTTCTGTGCAATTTGAAATTTGCCAGT
DB 136725 AACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6684 CGAGTGCCTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136665 TCACAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6744 AAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136605 GCAGAGAG-GAACAGTGTAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6804 AGCAGAAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
DB 136546 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6864 AGCCGAGATATCTGAGAGGTTTCTGTCAATTAATTAATTAATTAATTAATTAATTA
DB 136486 AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6924 CGATCATATGATCTTTACACTCTTTCTATGATGTAATGTAATTAATTAATTAATTT
DB 136426 AGAAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 6984 GAATTCCTTGACAGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136366 GGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 7044 TGTAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA
DB 136306 AGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 7104 CGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136246 AGCAGCAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 7164 AAATGCAAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136186 AGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 7224 AGCTTAGAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 136126 AATTAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 7284 CGAAGTTAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

```

Db	136066	AACAGCAACTGAGCAGCAGCAGCGTGGCGACGACGACGACAACACAGAGCAGCGATGAGC	136007
Oy	7344	AGCAGGTATTTCTACGAAGATATTCTTGTAAGCCGATAGAGAGATATCTCAGCGAGA	7403
Db	136006	AGCAGCACCATGAGCAGCAGATGAGCAAAAATATCACAGAAAGCAACAGAGCAGCGACGA	135947
Oy	7404	AGGAATTAAAGCATTTGTGAAGACTTCGATATTATTGGAAAAAATGTGATATTACAC	7463
Db	135946	AGAACGACGAGCAGCAGCAGCAGCACTGAGCAACAGCAGAGCAGCGACGACGCAAC	135887
Oy	7464	AGAGAGCAAAGATATATCATCTTCTACTGTGGTTTTGGGACTGACGGCTTGTCTCCGC	7523
Db	135886	CAGAGCAGCAGAAAGACGAGATATCAACAGAGAGCAGCAGCAGAGCAGTGTAAACGACGC	135827
Oy	7524	ATCAGGAA CAGTGCAGTTACAAATATTAAAGAAATTCGAGATTACTGTTAAAAATTC	7583
Db	135826	AGCAGCAGCAGCAGCAACGAGCAGCAGAGAGACAGCAGCAGGAAACGCAACAGCAGC	135767
Oy	7584	TTTTGTAAAGCAGCTGMAAAAGTAATGTTAGTCGATATTACAGAAATGTTGCCTTT	7643
Db	135766	AGCAGCAGCAGAGGTGTGCGCAGCAAGATAGCAGCAGCAGCAGCAGCGACCAGAGC---	135710
Oy	7644	AACAGCATATCAAGGTCTCTGTAGACATTTGGGAAATAGAGCTGCTTATGAGAAATTAA	7703
Db	135709	AGCAGAGCAGCAGAGGTCCAGCGACGACGACGACCAACAGAGCAGCAGCAACGAGAAAC	135650
Oy	7704	TTCTAATGGAAGATCAAAATATCATGATTATAAAATTTCTAAGCTATTAGAAAAATATTGA	7763
Db	135649	AGGAGAGCAGAGGAGAACTGTGTAAACAACAGCAGCAGCAACAGAGCAGCAGCAGCA	135590
Oy	7764	TGTTATTGTAAAAGTAAATTCGGAAATTGAGACCGGAAGCAAAAGATTAAACGTAGAGAC	7823
Db	135589	AGCAGCAGCAGAAAGCAGCAGCAAGAACAGCAGAGCAGCAGAGAGCAACTGAGAGCAGC	135530
Oy	7824	GGTAGCTGCCGAGCCATTATCTCAAAACCAAGTAATGAATTCAGAGTTGGAAT	7883
Db	135529	AGGTGACGACGACCAACGAGCAGCAGCATGACAGAGAGCAGAGCAGCAACGCAAA	135470
Oy	7884	TGAGAGAAGTATTTTCAATGAAGAAATAGAGTAAGTAAGTACCTCTTAAAGAAATTGGAAG	7943
Db	135469	CTGAGCAGCAGCAGAGGTGCGACAGGTGCACAGCAACAGAGCAGCAGAGCAGCAAC	135410
Oy	7944	AGAAATCAATGTCMAAGTGGAAAAAGAAACAGAGTGAATCTCAGAGACTTC	8003
Db	135409	AGCAACGTGAGCAGCAGCAGGAGCAACACAGCATGACGACGACGCAAAAATATCAAC	135350
Oy	8004	TGTAGAGCAGTAGCAGCGGGCAGGAATTATTTCCGAAGCAAAAAGATGCCGGAAGCTTTA	8063
Db	135349	AGGACCAACAGGAGCGACGACGACGACAGCAGCAGCAGGAGCGACGACGACAGAGCAGC	135290
Oy	8064	TTTGAAGTTAGTACAAAATCCGSAAGAGTATTTTTCATGCAATATGTGAATTGGA	8123
Db	135289	AGCAGCAACAGGAGCGACGACGACCAACAGAGCAGCAGCAACGAGAAAGCAACAGAGG	135230
Oy	8124	AGCAACACATAAAATGAAAGTAAACAGCAGTTTCTAAAGCAGTAAACAGGTTCTGTATTGGG	8183
Db	135229	AGCAGCAGGAGCGAGTGTAAACAACAG-----CAGGCGAAACGAGAAACACAGG	135176
Oy	8184	AGGAGTTGAGTCAACCAAGGCGAAGACTCTGCTGCAAGTAAACATATGTTAGAGTTGA	8243
Db	135175	AGCAGCAACATCAACTTGAAGCAGCAGAGGTGTGCGACGACGACCAACAGCAGCAGCAGC	135116
Oy	8244	GGNAGGAAATTTGTTCAAGAACAAAATGATTTGAATGCAATTTCTAAAGTAAAGTTTGA	8303
Db	135115	AGCAGAGAGCAGCAAGAAAGCAACAGAGCAGCAADAGCAGCAGCAGCAACACAACTTG	135056
Oy	8304	TGAAGATTAAAGTAACTGCTAAATCTTCTGTAGTATCAGGAAATGAGAGAGAAATTGCCGG	8363
Db	135055	AGCAGCAGCGCAGCAGGAAACAACAGAGCAGCAGGAGCAGCAGCAGCAACAGAGCAGC	134996
Oy	8364	AGCAGAGTGAATATCTTTACAGACACAA 8392	
Db	134995	AGGAGCGACGACGACGACGACCAAAAA 134967	

RESULT 27	AC153503/c	LOCUS	AC153503	203787 bp	DNA	linear	HTG 05-JUN-2005
DEFINITION	Mus musculus clone RP23-107G22,	WORKING DRAFT SEQUENCE, 6 unordered pieces.	AC153503	AC119819			
ACCESSION	AC153503.5	GI:66955894					
VERSION	HTG; HTGS PHAS1; HTGS DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFN.						
KEYWORDS	Mus musculus (house mouse)						
SOURCE	Mus musculus						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
REFERENCE	Muzny, D., Adams, C., Abbat, II, O., Allen, C., Alsbrooks, S., Archer, P., Arredondo, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R., Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P., Cadoree, J., Canada, A., Cardenas, V., Carter, K., Cavazos, I., Chacko, J., Chantour, M., Chavez, D., Chen, A., Chen, G., Chen, R., Cheng, M.-T., Chu, T., Clerc, K., Cockrell, R., Coyle, M., Cree, A., Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroil, L., De Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Donlin, J., McCanley, S., Dugan-Rocha, S., Dunn, A., Durbin, K., Dziuda, D., Egan, A., Escotto, M., Espinosa, Y., Eugene, C., Fa, M., Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P., Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T., Gaskin, C., Gench, S., Ghose, S., Gill, R., Gonzalez, D., Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlen, K., Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Havlak, P., Howe, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J., Hines, S., Hitchens, M., Hodgson, A., Hogues, M., Hollins, B., Howell, L. T., Hu, Y., Hume, J., Imo, K., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Kalatus, K., Kelly, S., Key, T., Khan, Z., King, L., Kovar, C., Kowis, A., Kowis, C., Lara, F., Leal, S., Lee, K., Lee, S., Legall, F. I., Lemons, S., Lewis, L., Li, B., Li, Y., Li, Z., Linell, M., Liu, W., Liu, Y.-S., Liu, Y., Liyanage, D., London, P., Lopez, J., Lorensheva, L., Lozada, R., Luk, T., Madu, R., Maheshwari, M., Mahoney, C., Mallory, K., Manouri, D., Martinez, E., McCelland, H., McPherson, J., Mercedado, C., Metcker, M., Milosavljevic, A., Minja, E., Morgan, M., Morris, S., Mundasa, M., Murray, D., Nazareth, L., Ngo, D., Nguyen, N., Norwig-Bastau, E., Nott, A., Nwokilemeh, O., Obregon, M., Ochi-Okorie, C., Odeh, E., Okunonu, G., Okunonu, K., Parker, D., Pasternak, S., Patel, B., Patel, V., Paul, H., Perez, A., Perez, L., Petrosino, J., Pham, T., Primus, E., Pu, L.-L., Puazo, M., Qin, X., Quinn, A., Quiroz, J., Rabat, D., Rachlin, E., Reigh, R., Ren, Y., Reuter, M., Richards, S., Rives, C., Rodriguez, F., Rojas, A., Ruiz, S. J., Sana, M., Sanders, W., Santibanez, J., Santos, R., Savery, G., Scherer, S., Shen, H., Shen, Y., Sisson, I., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sverek, A., Tamari, E., Taylor, T., Thomas, N., Thorm, R., Thornon, R., Trejos, Z., Usmani, K., Vargo, C., Verdunzo, D., Villasana, D., Virk, D., Vojkov, A., Waldron, L., Walker, B., Wang, Q., Wang, S., Warren, J., Wei, X., Wheeler, D., Williams, G., Williams, R., Worley, K., Wright, R., Wu, J., Yakub, S., Yan, K., Yuan, Y., Yu, F., Zhang, J., Zhang, L., Zhang, Z., Zhou, J., Weinstein, G. and Gibbs, R.						
TITLE	Direct Submission						
JOURNAL	Unpublished						
REFERENCE	2 (bases 1 to 203787)						
AUTHORS	Worley, K. C.						
TITLE	Direct Submission						
JOURNAL	Submitted (14-DEC-2004) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA						
REFERENCE	3 (bases 1 to 203787)						
AUTHORS	Worley, K. C.						
TITLE	Direct Submission						
JOURNAL	Submitted (05-JUN-2005) Human Genome Sequencing Center, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA						
COMMENT	On Jun 5, 2005 this sequence version replaced gi:66954712.						
	----- Genome Center						
	Center: Baylor College of Medicine						
	Center code: BCM						







Db 35552 AGCTGTGGCAGAGAGATGACGATCAGCACTTGCAGTAGTGTGGCAGAGAGTAGC 35611  
 QY AGTTACAAATATTTAAAGAAATTCGGAGTTACTGTGAAAATTCCTTTTGAAAGCACG 7598  
 Db 35612 AGCATCAGCAGATTGACGTGTGGCAGAGAGATGACGATCAGCAGTTGAAAGTAGC 35671  
 QY 7599 TGAAGAAATGTAATGTTAGTTCGATATTTACAGAAATGTTGCTTTAAGCAGATTCAGG 7658  
 Db 35672 TGTGGCAGAGAGATGACGAGCTGTGACAGTAGTGTGGCAGAGAGTAGCAGCAGCTT 35731  
 QY 7659 TCCCTGTAGAGAGATTGGGAATAGAGCTGCTATGCAAGATTAAATTTCAATGAAAGATC 7718  
 Db 35732 TGCAGTAGCTGTGGCAGAGAGATGACGAGCAGAG- TTGCAGTAGCTGTGGCAGAG 35790  
 QY 7719 AATATCAGATATTTAAATTTCTAAGCTATTAGGAAAAAATATTT---GATGTTATTGTA 7774  
 Db 35791 GAGTAGCAGCATCAGCAGATTGACAGTAGTGTGGCAGAGAGATGACAGCATCAGTAGTTG 35850  
 QY 7775 AAGATTAATCGGAATTGAGAGCGGAAGCAAAAGATTAAACGTAGAGCGGTAGCTCCG 7834  
 Db 35851 CAGTAGCTGTGGCAGAGAGATGACGAGCAGCAGCTTGCAGTAGGTGTGGCAGAGAGAG 35910  
 QY 7835 GAGCCATTATCTCAAAAGCAAAATGAATGAATGAAATTTAGAGTTGAAATTTGAAAGATA 7894  
 Db 35911 TAGCAGCAGTTGACAGTAGCTGTGGCAGAGAGATGACAGCAGTTGACAGTAGCTGTGGCAG 35970  
 QY 7895 TTTTCATGAGAAATAGAGTAAGTAAGCTAGCCCTTTAAAGAAATTTGAGAGAAATCAATG 7954  
 Db 35971 GAGGAGTAGCAGCAGATTGACAGTAGCTGTGGCAGAGAGATGACAGCATCAGCAGTTGCGAG 36030  
 QY 7955 TCAAAAGTGAAGAAAGAAACAGAGTACTGCTGAATCTCAGAGAGCTTCTTGAGAGCAG 8014  
 Db 36031 CAGCTGTGGCAGAGAGATGACGAGCATCAGCAGATTGACAGTAGCTGTGGCAGAGAGTAGCAG 36090  
 QY 8015 TAGCAGGGCAGGAATTTATTTCCGAGCAAAAGATCCGGAAGCTTTATTTGAAAGTTA 8074  
 Db 36091 CATCAGCAGTAGCAGTAGCTGTGGCAGAGAGATGACAGCATCAGAGTTGACAGTAGCTG 36150  
 QY 8075 GTACAAATTCGGAGAGATATTTTTCATGACAGATATGTAATGAAAGCAACATTA 8134  
 Db 36151 TGGCAGGAGAGATGACAGCATCAGCAGTTGACAGTAGCTGTGGCAGAGAGTAGCAGCAT 36210  
 QY 8135 AAATGAAGTAACAGCAGTTTCTAAAGCAGTAACAGCTTCTGATTTGGAGAGAGTTGGAG 8194  
 Db 36211 CAGCAGTTGACAGCAGCTGTGGCAGAGAGATGACAGCATCAGCAGTTGACAGCAGCTGTG 36270  
 QY 8195 TCACCAAGCAGAGAGTACTGCTGACAGTTAAACATATGTAAGTTGAGGAAGAAATT 8254  
 Db 36271 CAGGAGTAGCAGCATCAGCAGTTGACAGCAGCTGTGGCAGAGAGATGACAGCATCAGCAG 36330  
 QY 8255 TGTTCAGAACAATCAGATTGAATGCAATTTCTAAAGTAGAAGGTTTGATGAAGATTAAG 8314  
 Db 36331 TTGCAGTAGCTGTGGCAGAGAGATGACAGCATCAGCAGTTGACAGTAGCTGTGGCAGAG 36390  
 QY 8315 TAACTGTAATTTCTGTATGATCAGAGAAATGAGAGAGAAATGCGGAGCAGAGTGA 8374  
 Db 36391 GAGTAGCAGCATCAGTAGTTGACAGTAGCTGTGGCAGAGAGATGAGAGCAGCAGCTGTG 36450  
 QY 8375 ATACTCTACAGCAAAAGTAATATGAAATCCGTAAGTCTTATACAAAGCAAGATTATG 8434  
 Db 36451 CAGTAGGTGTGGCAGAGAGATGACAGCAGCAGCAGTAGTGTGGCAGAGAGAGAG 36510  
 QY 8435 AAATATATGATTAACAAAAAATATATTTTACAGAGTCAATGCTCTTTTAATGATA 8494  
 Db 36511 TACAGAGAGCAGAGATTGACAGTAGCTGTGGCAGAGATGACAGCATCAGCAGTTGCGAG 36570  
 QY 8495 CAAGAATGAAGCAATATATGAATCTTTAGCGGTAGCCGCTGTGCATGCACAAAGAAACA 8554  
 Db 36571 CTGTGGCAGAGAGATGACAGCATCAGCAGTTGACAGCAGCTGTGGCAGAGAGAGTAGCAG 36630  
 QY 8555 ACAAGCATTTAGAGAG-----TCAAACAAGTTAATTTTACAACTGTAATGAGAGAA 8608

Db 36631 CATCGCAGTTGACAGCAGCTGTGGCAGAGAGATGACAGCATCAGCAGTTGACAGCAGCTG 36690  
 QY 8609 ACGTATCTCAACTTGTGCAAAAGCTTTGGCTAAATATGAAATATATGAAATGTAAG 8668  
 Db 36691 TGGCAGAGAGATGACAGCATCAGCAGTTGACAGTAGCTGTGGCAGAGATGACAGCAG 36750  
 QY 8669 GAACTGAGAGAGCTTTAGTTCGAGAGCAAAAGCAGCAGCCGTTGAAATTAATCAAGAGTA 8728  
 Db 36751 CAGTTGAGTAGCTGTGGCAGAGAGATGACAGCAGCAGCAGTAGTTGACAGTAGCTGTGAGAG 36810  
 QY 8729 CTACAGAGCATTGTTGACAGAAATTTGGAAATTTGAGATTAATTAAGAAAGATTGCA 8788  
 Db 36811 TAGCAGCATCAGCAGATTGACAGTAGCTGTGGCAGAGAGATGACAGCATCAGCAGTTGCGAG 36870  
 QY 8789 GAGATTAATCACTTTAAGATGCAACGAGACGGAACCAAGAGAGAGTCTGTGCGAAAGA 8848  
 Db 36871 TAGCTGTGGCAGAGAGATGACAGCATCAGCAGTTGACAGTCTGCTGTGGCAGAGAGTAG 36930  
 QY 8849 ATGTATTTTGTGTAATAATCAATTTTCAGGGGAAACAAATCATTCATTGAAGATAAG 8908  
 Db 36931 CAGCAGCAGCAGTTGACAGTAGCTGTGGCAGAGATGACAGCATCAGTAGTTGACAGTAGCTG 36990  
 QY 8909 CCGAATTTGTTGAACCGGAAGTGAATGTAATGATGCTTTGAATGAATTGATGATG 8968  
 Db 36991 TGGCAGAGAGATGACAGCATCAGCAGTTGACAGTAGCTGTGGCAGAGAGATGACAGCAT 37050  
 QY 8969 TACAAAGAAAGTAGTGTGCTATGTTGGAATTTGTAATGGAATGTAATGTA 9021  
 Db 37051 CAGCAGTTGACAGTAGCTGTGGCAGAGAGATGACAGCATCAGCAGTTGACAGTA 37103

RESULT 29  
 AC122227  
 LOCUS 208332 bp DNA linear ROD 01-JAN-2004  
 DEFINITION Mus musculus chromosome 12 clone RP23-131H10, complete sequence.  
 AC122227  
 VERSION AC122227.4 GI:40539152  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 208332)  
 Wilson,R.K.  
 The sequence of Mus musculus clone  
 Unpublished  
 2 (bases 1 to 208332)  
 McPherson,J.D. and Waterston,R.H.  
 Direct Submission  
 Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 3 (bases 1 to 208332)  
 McPherson,J.D. and Waterston,R.H.  
 Direct Submission  
 Submitted (28-FEB-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 4 (bases 1 to 208332)  
 Wilson,R.K.  
 Direct Submission  
 Submitted (01-JAN-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 On Jan 1, 2004 this sequence version replaced gi:26604203.

COMMENT  
 ----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu  
 Contact: submissions@wustl.edu  
 ----- Project Information -----  
 Center project name: M BA0131H10  
 -----  
 Location/Qualifiers

source 1..208332  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="12"  
 /clone="RP23-131H10"

Query Match 1.1%; Score 107.4; DB 9; Length 208332;  
 Best Local Similarity 41.8%; Pred. No. 1.8e-05;  
 Matches 722; Conservative 0; Mismatches 1006; Indels 1; Gaps 1;

QY 7105 GTACGAGAAAAAGTAATGTAAACGCAAAAGAGAAAAAATATTAAAGCAACAGCAGCA 7164  
 DB 106695 GAACAAGAGAGAACAG 106754  
 QY 7165 AATGCAGGAATCGGAGAGACGACATCGAGCCATGTCTTGTAATATTTTGGACAA 7224  
 DB 106755 GAAGAAAG 106814  
 QY 7225 GCTGTAGAGATAGAAAAAATCTGAAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAG 7284  
 DB 106815 GAAG 106874  
 QY 7285 GAAGTTAACAAAGAACAGATAAAAAGTAAATGATCTGCAAAAAAATCTTACATCA 7344  
 DB 106875 GAAGAAAG 106934  
 QY 7345 GCAAGTATTTCTACAGAGATCTCTGTAAAGCGGATAGAGAGATCTCAGGAGAA 7404  
 DB 106935 GAAG 106994  
 QY 7405 GGAATTTAAAGCATTGTGAAGACTTCTGATTTATTGAAAAAATGTAGATTTACA 7464  
 DB 106995 GAAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 107054  
 QY 7465 GAAGACAGAAATTAATCACTTCTACTGTGTTGGGAATGCAAGCTTCTCCGCA 7524  
 DB 107055 GAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 107114  
 QY 7525 TCAGGACAGTGGCCTTCAATATTTAAAGAAATTCGAGGATTAATCTTGAATTTCT 7584  
 DB 107115 GAAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 107174  
 QY 7585 TTTTGCAAGACAGCTGAAAAAGTAAATGTTAGATCGATATTACAGGAATGTTGCTTA 7644  
 DB 107175 GAAG 107234  
 QY 7645 ACAGCATATCAAGTCTCTGTAGAGCATTGGAAATAGAGCTGCTATGCAATTAAT 7704  
 DB 107235 GAAG 107294  
 QY 7705 TCTAATGAGAGATCAATATTCAGTATTTAAAAATTTCTAAGCTATTAGAAAAATTTGAT 7764  
 DB 107295 GAAG 107354  
 QY 7765 GTTATTGTAAAGATTAATCGAATTTGAGAGCGAGAGAGAGAGAGAGAGAGAGAG 7824  
 DB 107355 GAACAAG 107414  
 QY 7825 GTAGCTCCGAGCCATTATCTCAAAAGCAAGATGAATTCAGAGGTTGAATTT 7884  
 DB 107415 GAAG 107474  
 QY 7885 GAGAGAGATTTTCAATGAAGAAATAGATTAATGAGCTTCTTAAAGAAATTTGAGAG 7944  
 DB 107475 AAG 107533  
 QY 7945 GAATCAATGTCAAGTGAAGAAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 8004  
 DB 107534 GAACAAG 107593  
 QY 8005 GTAGAGAGATGAG 8064

DB 107594 GAAG 107653  
 QY 8065 TTGAAAGTATGATCAAAATCCGAGAGAGATTTTCTTGCAGATAATGTGATTTGAA 8124  
 DB 107654 CAAG 107713  
 QY 8125 GCAACATTAATTAATGAAGTAACAGCAGATTTCTAAGCAGTAACAGTTCTGTATTGGGA 8184  
 DB 107714 CAAG 107773  
 QY 8185 GAGTTGAGATCACCAG 8244  
 DB 107774 GAAG 107833  
 QY 8245 GAAG 8304  
 DB 107834 GAAG 107893  
 QY 8305 GAAGATTAAGTAACTGTAAATCTTCTGTAGATCAGAGAAATGAGAGAGAGAGAGAG 8364  
 DB 107894 GAAG 107953  
 QY 8365 GCAAGAGTGAATTAATCTTACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8424  
 DB 107954 CAAG 108013  
 QY 8425 CAAGTTTGAAGATTAATGATTTACACAAAAAATATTATTGAGAGTCAATGCTTTGCT 8484  
 DB 108014 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 108073  
 QY 8485 TTAATGATTAACAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8544  
 DB 108074 GAAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 108133  
 QY 8545 CAAG 8604  
 DB 108134 CAAG 108193  
 QY 8605 GGAACGTATCTCAACTTCGTGCAAAAGCTTTGCTTAATTAATGAATTTAGAGAA 8664  
 DB 108194 CAAG 108253  
 QY 8665 AAAG 8724  
 DB 108254 CAAG 108313  
 QY 8725 AGTACTACAGAGCAATGTTGTTGACAGAAATTTGGAAATTTGAGATTAATTTAGAA 8784  
 DB 108314 GAAG 108373  
 QY 8785 GCAAGAGATTAATGATTTGTAAGACTCAACGAGAGAGAGAGAGAGAGAGAGAGAG 8843  
 DB 108374 GAAG 108422

RESULT 30  
 AC116998/c 216959 bp DNA linear ROD 08-NOV-2003  
 LOCUS AC116998 Mus musculus BAC clone RP23-16f6 from 10, complete sequence.  
 DEFINITION AC116998  
 ACCESSION AC116998.3 GI:24080765  
 VERSION  
 KEYWORDS HTG.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridae; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 216959)  
 AUTHORS Isek, A., Kozlowicz, A., Spalding, L., and Mangiapanello, L.  
 TITLE The sequence of Mus musculus BAC clone RP23-16f6  
 JOURNAL Unpublished (2001)  
 REFERENCE 2 (bases 1 to 216959)

AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Wilson, R.  
Sequencing of Mus musculus  
Unpublished (2001)  
3 (bases 1 to 216959)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (05-APR-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 216959)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (28-MAY-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 216959)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (17-OCT-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
6 (bases 1 to 216959)  
Wilson, R.  
Direct Submission  
Submitted (08-NOV-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Oct 17, 2003 this sequence version replaced gi:21218578.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [submissions@wustl.wustl.edu](mailto:submissions@wustl.wustl.edu)  
----- Summary Statistics  
Center project name: M\_BA0016F06  
-----

NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren,  
Department of Genetics, Washington University, St. Louis MO. For  
additional information about the map position of this sequence, see  
<http://genome.wustl.edu>

SOURCE INFORMATION:  
The RPC1-23 BAC Library has been constructed by Kazutoyo Osegawa  
and Minako Tateo in the laboratory of Pieter de Jong  
(<http://www.chori.org>) from female C57BL/6J mouse kidney and/or  
brain genomic DNA. The clone and detailed information can be  
obtained from Research Genetics, Inc. (<http://www.resgen.com>) or  
Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:  
This sequence is the entire insert of the clone.  
Location/Qualifiers  
1. 216959  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="18"  
/map="18"  
/clone="RP23-16Fc"  
/clone\_1id="RPC1-23"  
1271..1406  
/rpt\_family="BRVK"  
repeat\_region

repeat\_region 1423..1535  
/rpt\_family="B4"  
repeat\_region 1611..1806  
/rpt\_family="RMERIA"  
repeat\_region 3133..3210  
/rpt\_family="ERV1"  
repeat\_region 3256..3331  
/rpt\_family="ERV1"  
repeat\_region 3343..3424  
/rpt\_family="ERV1"  
repeat\_region 3450..3501  
/rpt\_family="ERV1"  
repeat\_region 5895..6220  
/rpt\_family="MALR"  
repeat\_region 6891..6948  
/rpt\_family="MALR"  
repeat\_region 12422..12679  
/rpt\_family="B4"  
repeat\_region 13130..13212  
/rpt\_family="ERV1"  
repeat\_region 14149..14202  
/rpt\_family="Alu"  
repeat\_region 15684..15797  
/rpt\_family="MIR"  
repeat\_region 15937..16017  
/rpt\_family="Alu"  
repeat\_region 16859..17015  
/rpt\_family="Alu"  
repeat\_region 17017..17905  
/rpt\_family="BRVK"  
repeat\_region 17924..18007  
/rpt\_family="Alu"  
repeat\_region 18237..18398  
/rpt\_family="MIR"  
repeat\_region 18742..18905  
/rpt\_family="L2"  
repeat\_region 19092..19510  
/rpt\_family="MALR"  
repeat\_region 20001..20444  
/rpt\_family="MALR"  
repeat\_region 20944..21019  
/rpt\_family="ID"  
repeat\_region 21960..22289  
/rpt\_family="BRVK"  
repeat\_region 22439..22792  
/rpt\_family="MALR"  
repeat\_region 23315..23458  
/rpt\_family="Alu"  
repeat\_region 23821..23968  
/rpt\_family="Alu"  
repeat\_region 23985..24243  
/rpt\_family="L1"  
repeat\_region 24258..24469  
/rpt\_family="B2"  
repeat\_region 25483..25910  
/rpt\_family="L1"  
repeat\_region 26435..26691  
/rpt\_family="B4"  
repeat\_region 26763..26867  
/rpt\_family="L1"  
repeat\_region 27793..27942  
/rpt\_family="L1"  
repeat\_region 28085..28193  
/rpt\_family="Alu"  
repeat\_region 28584..28656  
/rpt\_family="Alu"  
repeat\_region 28814..28948  
/rpt\_family="Alu"  
repeat\_region 29355..29568  
/rpt\_family="B2"  
repeat\_region 29717..29907  
/rpt\_family="B2"  
repeat\_region 30233..30313

[illegible]

Db	37523	AGCAGTAGCAGTACGACGAAAGCTGTAGCAGTAGACAGTAGACAGTAGACAGT	374644
Oy	6795	AGATGTAATATGACGAAATAGATTAATAATCAGGTACTGGAAATCCGTTCCAGCCGAGCTGG	6854
Db	37463	AGCAGTAGCAGTAGCAGTACCTAGCAGTAGACAGTAGACGAAAGCTGTAGCAGTAGACAGT	37404
Oy	6855	AATTCCTTGACCCCGAGTATCTGAGAGTGTCTCTGTCAATTAATATTGCAAAATTAAGTAGA	6914
Db	37403	AGCAGTAGCAGTAGCAGTAAGAACTGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAGT	37344
Oy	6915	AACAGATATCGATCATAGTACTTTACACTCTTTACTGATGTGTAAAGTAAAGCTCTTAA	6974
Db	37343	AGCAGTAGCAGTAGCAGTAGCAGTAGACAGTAGACAGTAGACAGTAGCAGTAGCAGTAGACAGT	37284
Oy	6975	TAAATTTTCGAATTCCTTGACAGCCCGTGGAGAGCCGAGGCTTGACAGCAGTTAACCG	7034
Db	37283	AGCAGAAAGCTGTAGCAGTACGACAGTAGACAGTAGACAGTAGACGAAAGCTGTAGCAGTAGACAGT	37224
Oy	7035	AGTGGTTTTCTGTTAACAATAATAGTTCTGTGATAGCTCGAGTTTCAACAATPACTTGA	7094
Db	37223	AGCAGTAGCAGTAGCAGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAG	37164
Oy	7095	TTTGACTTCCGTACGAGAAATAATGTAAATGTAAAGCCAAAGAGAAATAATTAAGCA	7154
Db	37163	AGCTGTAGCAGTAGCAGTACGACAAAGCTGTAGACAGTAGACAGTAGCAGTAGCAGTAGACAGT	37104
Oy	7155	AACAGCAGCAAAATGACGAAATCGAGAGACAGCAATCCGAGCCAAATGTCTTGAGTAAATPA	7214
Db	37103	AGCAGTAGCAGTAGCAGTAAGAACTGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGT	37044
Oy	7215	TTTTGGAACAGCTGTAGAAAGATAGAAAAATTTGGAAGAAAAAGAACAAAGTTTAA	7274
Db	37043	AGCAGTAGCAGTAGCAGTAGCAGTAGACAGTAAGAAAGCTGTAGCAGTAGACAGTAGACAGT	36984
Oy	7275	AACCTTAGACGAA--GTTAACAAGAAACAAGATTAATAAAGTAAATGATGTCTAACAAAA	7331
Db	36983	AGCAGTAGCAGAAAGCTGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAGT	36924
Oy	7332	AATCTTAACAATCAGCAGGTATTTCTACAGAAAGATCTTCTGTAAAGCCGATAGAGAGA	7391
Db	36923	AGCAGTAGCAGTAGCAGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAGT	36864
Oy	7392	TACTCAGGAGAAAGAAATTAAGCACTTGTGAAGACTCTTGATATTATTGAAAAAATGT	7451
Db	36863	AGCAGTAGCAGTAGCAGTAGCAGTAAGAAAGCTGTAGACAGTAGACAGTAGACAGTAGACAG	36804
Oy	7452	AGATATTACAAGAGACAAAGATTAATCACTTCACTGGTGTGGAACTGCAG	7511
Db	36803	AGCTGTAGCAGTAGCAGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAGT	36744
Oy	7512	TCTTGCTCCGCATCAGAAACAGTGGCAGTTAACAAATATTAAAGAAATTCGAGATTAC	7571
Db	36743	AGCAGTAGCAGTAGCAGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAGT	36684
Oy	7572	TGTTGAAATTCCTTTGTGAAAGCAGCTGAAAAAGTAAATGTTAGATCCGATATTACAG	7631
Db	36683	AGCAGTAGCAGAAAGCTGTAGCAAGTAGACAGTAGACAGTAGACAGTAGACAGT	36624
Oy	7632	AAATGTGCTTTAACAGCATATCAAGTCTGTAGAGACATTGGAAATAGAGCTGCCTA	7691
Db	36623	AGCAGTAGCAGTAGCAGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAG	36564
Oy	7692	TGCGAAATTAATTTCTAATGGAAGATCAAAATATCAGATTTAAAAAATTTAAGCTATTAG	7751
Db	36563	AGCTGTAGCAGTAGCAGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAG	36504
Oy	7752	AAAAAATATTGATGTATTGTAATAAGATTAATCCGAATTGAGAGCCGAAAGCAAAAGATT	7811
Db	36503	AGCTGTAGCAGTAGCAGTAGCAGTAGACAGTAGACAGTAGACAGTAGACAGTAGACAGT	36444
Oy	7812	AACGTAGAGACGGTAGCTGCCGAGCAATTATCTCAAAAGCA	7854



EDKALNEKIMSEEEYVIEDONHEVPAEVSUTEI PGVEETTNESHVKGENVEN  
VVDNSVDEIQFDDDSIEITYDSKDFHKEENYDSFREERDENIHSRKNN  
FVKIENKSVGVHPTTEYSKENIAEVPVSHIMPEKVENPEVPEVSEIIEVE  
BEVTSKENIEEAPGDINSENIIEEVPVSEIIVONDNINTNTIDHFDLD  
DLKIPEAKRKLKRLTYDVLNIDVIEKQRBEMKKNRDFYICLRINIKSEPTLM  
KEYGSDSNKSANDNTYVYIKLIKQKQGRIVDTIANSLSYSLNLYIEMEKD  
LILNKTETNKSQKHNKUDMKWKEKMFELFKDLKIDMKRTKISIDQNEENESM  
NLSDELKNKTSDMNRNEQKLEKNEKSNLSKDLQIEKQIIMKSNIVKVNIE  
NMFDEKQVNVNELEKNDVSNLOENSEASVENYISDVDSQDMNTSINDESK  
IVASSKTNIEEYIILSRKDLINIIIVMHVMDLPQKIDELKAKKFLNLSIDFI  
KEKIKDKEVLDYIIDIIRIIVEDTSHVNNIKKAIIDHTSHDMFRLRQNAON  
SISQVDILKEKRYKILISBEDKKNENNNDVEVPELKDNKCEBNSFNLRSNK  
KQDPEEKKNKNDLNTLISDSKININIEIIESKNGENDKNSGNNSRKRDR  
RTEREKQDLRVOLQDYENIFEOIOMNENONKQKNETYNIKTSIDDKNLLRE  
YKKEKNIINOSKEFSVNNEN"

## ORIGIN

Query Match 1.1%; Score 105.6; DB 2; Length 16377;  
Best Local Similarity 41.8%; Pred. No. 7.9e-05;  
Matches 1838; Conservative 0; Mismatches 2479; Indels 75; Gaps 17;

QY 5131 GATGAAACGGAAGCTTAGTGTGATCTGAGTTTGAAGAGTAACTTTCATGTA 5190  
DB 3325 GTTGTGAAAAAGATCAGATCTGAAATTTGTTGTAAGAAAGATCAGCTAGTAAAT 3384  
QY 5191 GATGCAAAAGATCAAAAAACAATAATGACGGAATG---CAATGAGAGAAA 5247  
DB 3385 GTTGTGAAAGAGATCAGCTACTGAAAGTGTGTGAAAGAGATCAGCTTCTGAAAG 3444  
QY 5248 GGGGCTGGAGTTGAGCAACAGTTGCTACCAATATTGGAACAACTAGTTAGT 5307  
DB 3445 ATGTGAAAGAGAGATCAGCTACTGAAAGAGTGTGAAAGAGATTAATCTAGTAT 3504  
QY 5308 ATTGTAAAAACAGTAAATTTACAACGCGGAATGATCAGATGAAAAAATATCATGTG 5367  
DB 3505 AATGTTCAAAAGTCAAAAAGAGATTAATTGAAAAAGTAAATATCTGTAAABACA 3564  
QY 5368 ACTGCAAAAGATTATCTATGACCAATCTATAGCAGTCGAGATTGAGAGCAAAAAGA 5427  
DB 3565 GCTAAGAGATGAAATCTATGATGAAAAAATACCATGAGAAATCATCTGTTGTATGAT 3624  
QY 5428 GCTCTGTGCAAGAGAGCTTCTGCAAGTACTACCTGGAATTAAGACGTTTCTCATGTT 5487  
DB 3625 GAGTCTTTTAAAGACAAGAGACCTACAGATTAATGATCAAGACGTAATTCGGA 3684  
QY 5488 GATCAAACTGATATTGACAAAGATTAGAGAAAGAAATATGAGAAATAGAGAAAAGCA 5547  
DB 3685 GATATTATTATTAAGACATCTCCATTAGAAAGAACAAAATTGAAATTTACCAACAGA 3744  
QY 5548 AATGTTAATGTTCTAGCTGAAAAATACGA---GTCAAGTGTCACAATGCGACAGTCTT 5604  
DB 3745 TATATTACTACTGCTGATATACACTAGAGGTGAAACAGAAACAAAATATATCTAATA 3804  
QY 5605 TCCGAGCAAGTGAACAAGCTGAGTGAAGCTGAGTACGATTATATAATACACA 5664  
DB 3805 TACGAAAAATTAATGAAAGAGTGAAGAAAGCAAAATTTCAAGAGAAAAATTAACGA 3864  
QY 5665 AATPACTTGCACATATAAAAAATAGTACTCAAAATGTACGAATGCTTGTAAAAAAC 5724  
DB 3865 AATATTCCAGTAGAAAGAGATCAGTTACAGAAAGATATAGTACAAAGACCATCATTAGCT 3924  
QY 5725 AATCTCATTCATCTAATTAACAATTTGGAATTTGAGCTGAGTTGAGCTGAGAGAGCT 5784  
DB 3925 CAGGAAGTTGAACAAAAGATCAGATCTAATGAAATTTGAGAAACATAATTAGCTAT 3984  
QY 5785 GGAAGTACAGGTTCTAGACAGTGAATAAGATTGTAATATACATAGACAGATTAAAT 5844  
DB 3985 GAAAAAATATATCCAGAGATCTGTTACTGAAAAAGTGTGAAAAAAGAGATAGAT 4044  
QY 5845 CATGCAAAATCATCTGCAAGGGAATGTGAGATTATTAAGAGCTGATCGGTAAT 5904  
DB 4045 ACTGAAGAGGCTTTAGAGAGATGATCAATTACTGAAGAAATAGTAGAAGAGAGTA 4104

QY 5905 GCTAATTATGACAGAACGCTGTGAGAGTGCCTGCGAGCAATAGAGCTCAACAGT 5964  
DB 4105 TCAGCTTGTGAAGAAATAGTAGAGAGAGAAATCATCTAGTGAAGAAATAGTAGAGAA 4164  
QY 5965 GTGAATGAATAATACAGATCTACAAAGCATATGTAAAGATTCTA---CAGTATGCT 6021  
DB 4165 GAGATCATCTAGTGAAGAAATAGTAGAGAGAGAAATCATCTAGTGAAGAAATAGTA 4224  
QY 6022 AAAGAGAAACAGATGATTAATTAATCTACTCAAGGCAAGTAGATTAAGGTAGATAAA 6081  
DB 4225 GAGAGAGAAAGATCATCTAGTGAAGAAATAGTAGAGAGAGAAATCATCTAGTAGAGAA 4284  
QY 6082 GTATTCAAAAATCTTAATATTTAGCAGAGACTTATCACAAAAAAGAAAAATAGTATAA 6141  
DB 4285 ATATAGAGAGAGGTCAAGTACTGAAAGAGTCTTAGAAGAGAGAGAGATCGGTTACGGA 4344  
QY 6142 AAAAGATTTGTTCACATAGTTCACTCTCATCTTAATTAATCTTTATTGGCAAAATCC 6201  
DB 4345 GAAATTTGAGAGAGAGAGATCACTACTGAAAGATTAAGATATTGATCAGTTAGC 4404  
QY 6202 GCTGGTTCAAGACAAAGCCGAGTGCAGAGAACTGTTAATTAACAACAGTTTATGCA--- 6258  
DB 4405 GAAAGAGTATTAGAGAGAGAGAGATTCGCTACTGAAAGATTTGTAGACACACAGATCT 4464  
QY 6259 GAAACAGAGCTCTTTGAGAAATTTCTATATTAAATGCAAAACATTATCTGTAAATCA 6318  
DB 4465 GTTATTGAAGAAATGTGAGAAACAGAAATCATCTACTGAGAAAGTTCTAGAAAGTTGGA 4524  
QY 6319 GAGATTTACAGAAATTCATGCGAGTACTAGTTCTGTGTGTG---GTGGAATGTA 6375  
DB 4525 TCAATGTGGAAGAAATGTACAGAGAGAGGCTGTGTGCTCAAGAAATGTACATGAA 4584  
QY 6376 GAGAGAGAGCTTCTTCTGATACCAATTTATTAAGAAATATCCAGACAAAGATTGGA 6435  
DB 4585 GAAATATCACTACTGAAAAACATGATGAGTAAATGATCACTACGAGAAATTTGTA 4644  
QY 6436 AAAC-----TACAATGTGATGAGAGTTTCGAGAGAGAGCTGAAATTTACAGAGAT 6489  
DB 4645 GAAAAAGTTGATCAGTTTCTGAGAAATTAATAGTAAAGAGAGTCACTTCTGAAAGAA 4704  
QY 6480 TCTAAGCAAGAAATTTCTCTTTTGGAGTCCGAGTCGACAGACCCGGGTAGAGCCGA 6549  
DB 4705 ATAGTAGAAGAGATCAGTTACTGAAAGATTTGAGAGAGAGAAATTAATTAATGAA 4764  
QY 6550 GTGCAAGAACGTTCCGTAAATCAATTTGACAGAAAGACGAGATGATGTGAAAGAA 6609  
DB 4765 GTTGTGAAACCGAATCGGTACTGAGAAATATGACAAAAAGAGATACAGATGCTGAA 4824  
QY 6610 GCAAGATTTTGTAAAA-----AAAGCTGATTAACGCAAAACGTTATAGTTCT 6660  
DB 4825 GAAATTTTAGACAAAGAGATCAATGATGAAAGAAATATTAGAAAAAGAAATCAATTGTT 4884  
QY 6661 GTTGCATTTGAAATGCGCAGTCGAGTGGCTGCAAAAGAGCTGGAATTTGAGACA--- 6717  
DB 4885 GAAAGAAATTTGTGACCGAGAGAGATCTGTTCTGAAAGAAATAGTAAATTAACGATCAATT 4944  
QY 6718 GCAGTGGCAGTTTACCAAGATGAAATCAACAGAGCAAGAGTGAATAATTTCTAAAAAT 6777  
DB 4945 GCTGAAGAGTCAAGAGAGAGAAATTAAGTTACAGAGAGAGCTGACATGATAGAGATCT 5004  
QY 6778 ATGACTCGAAAACAATTAGATGTAATAGCAGAAAAAGAGATTAATTCAGAGTACGGAATC 6837  
DB 5005 GTTACTGAAGAAATTAAGAGAGAGAAATCGATTACTGAGAAATGAAAGCAATAGAAATCA 5064  
QY 6838 GGTTCAGCCGAGCTGGAATTTCTGACACCGAGATCTGAGAGTGTGTTCTGTCAATAT 6897  
DB 5065 GCTTTTGCAAAATTAATGAGAGAAAGAGAACCAATTAATGCGAAATTTGTAAGAGAGAA 5124  
QY 6898 ATTGCAATTAAGTGAAGACATATGATCATAGTACTTTACTTTCTACTAGTGA 6957  
DB 5125 GATTAAGTACTGAGAGAAATAGTGAATGAAATATAGTTACTGACGAGAGCTATAGAGAA 5184  
QY 6958 AATGTAAAAAGCTTTAATTAATTTGAAATTTCTTGAACGCGGTGAGAGAGCGCAGGT 7017

Db	5185	GAATAATTAGTTAATGAACTTTGATGTAAGAAAGATCTGTCTACTGAAAAACAGTA	5244
Qy	7018	CTTCACAGACTTACCGAGAGTGTTTCGTTAACACTAAATAGTCTGTAGATGCTGA	7077
Db	5245	GAAGTTGAACGATCAGCTACTGAAGAATCTGTGGAGAAAGACATCAGTTACTGAAAA	5304
Qy	7078	GTTCCAAATACTCTGAATTTGACTTCCGTACGAGAAAAATGAATGTAAACGGCAAAAG	7137
Db	5305	GTTTCAGTGCAATGAAGATCAACTACTGAAC---AAATACTAGATGAATCAGTTGGTGA	5361
Qy	7138	GAATAAAATATTAACCAACAGCAACAAATGCAGAAATCGAGAGACGCAATCGAGCC	7197
Db	5362	GAATATGATAGAGAAAGATCTGTGATGACAAAATAATAGAGGAAGATATCTGTT	5421
Qy	7198	AATGCTCTGTGAATTAATTTTGGAAACGCTGTGAAAGATGAAAAATTCGAAGGAAA	7257
Db	5422	GATGAAGTTGTAGAAAGAAAGATCTGTTATTAAGAAATAGTAGAAGAAAGAAATCA	5481
Qy	7258	GGAAACAGAGTTTAAAACTTTAGACGAAGTTAAACAAGAACACAGATTAATAAAAGTAAT	7317
Db	5482	GTTCTGAGAAATTTTGGAGAAAGAACTATCAGGTTCTGAAGAAAGTTCTAGAAAGCA	5541
Qy	7318	GATGCTACGAAAAAAATCTTAACAATCAGACGTAATTTCTACAGAAATCTTCTGTA	7377
Db	5542	TGGGTTACTGAATGCTTTATGAGGACAAAGAGATCTGTATTAGAAAGAAATTTGA	5601
Qy	7378	GCGGATAGAGATATCTCAGAGAGAAAGATTAAAGCATTTGTGAAGCTTCGTGATAT	7437
Db	5602	GTAGATGAGAAAGATTCATTAACGAAATAATGTAGAAAGTGAATCAGCTAATGAAAA	5661
Qy	7438	ATTGAAAAAATGTGATATTACAACAGAGGCAAGAAATATATCATCTTACTCGTG	7497
Db	5662	ATAGTAGAGAGAAACATCTCTGTGTAAAGATTTTAGAAAAAGAAATTTGTTAT	5721
Qy	7498	TTGGAACTGCAGGCTCTTGCTTCCGCATCAGAAACAGTGCAGTTACAAATATTTAA	7557
Db	5722	GAAAGAAATATAGAAAGAAAGGTCGTAAATTGAACAAGTTGAAGATAC-----	5775
Qy	7558	AAATCCGAGTTACTGTTGAAATTTCTTTGTGAAGACGCTGAAAAAGTAAATGTT	7617
Db	5776	GTTTCTGAAAAATCAGAAAGATCATCACTATTGAAGAAATTTAAAGAAATTAAG	5835
Qy	7618	TGCGAATTAACAGAAATGTTGCTTTACAGCATATCAAGGCTCTGTAGAGCATTTGG	7677
Db	5836	GAATCGATTAGAAAAAATAGTAAGAAAGAAATCTGTTACTGAAGAAATGTACGA	5895
Qy	7678	ATAGAGCTGCCTATGCAGAAATTAATTTCTATGAGAGATCAAAATATCAGTATTA	7737
Db	5896	CAAGAGAAATCTACTACTGAAAAAATAGTAAAGATGTATCA-----CTACTGA	5949
Qy	7738	TCTAAGCTTATAGAAAAAATATGATGTTATTTGTAAGATTAATCGAAATTGACGC	7797
Db	5950	TTTGATGAAACAAACAGATTCAGTTACTGAAAAAGTTATAGAACAAAGAGATCA	6009
Qy	7798	GAAAGAAAGATTAACCTAGAGCGGATGTCGCGAGCATATATCTCAAAAGCAAG	7857
Db	6010	GAAAGTTCAGAAAGCCTTAAGAAAAAGATCCGTTCTGATGAACATGAACAGAGAT	6069
Qy	7858	AATGAATGAATTCACAGGTTGAAATTGAGAGAGTATTTTCATGAGAA--AATAG	7915
Db	6070	GTTACTGTAATGCAAGATTAATGTGAGAAAAATCAGTTACTAAAGAAATGTGAC	6129
Qy	7916	TAACTAGCCCTTCTAAGAAATTTGAGAGAA-----TCAATGTCAAGTGA	7965
Db	6130	GAGTGCTCTGACTGAAGAAATAGTAGAAGAAATTCATCAGTTACGAAAGAGTAC	6189
Qy	7966	AAAGAAACAGATGCTGTAATCTCAAGAGCTTCTGTAGAGCAGTATGACAGGCA	8025
Db	6190	GAAACAGGATCCGTTACTGAAGAAATATAGAAAGCAACAGTAACTGAAGAAATCG	6249
Qy	8026	GGAATATTTCCAAACAAAGATGCGCGAAGCTTATTTGAAAGTTAGTCAAAATCC	8085

Db	6250	GAAGATGCATCAGTAACTGTAAGAAGTCGTAAAGATTCAGTTATTCAGAAAGTTGTA	6309
Qy	8086	GGAAGAAGTATTTTTCATGCAGATTAATTCGATATATGAAAGCAACACATATAATGAAAGTA	8145
Db	6310	GAAGATGCATCAGTAACTGTAAGAAGTCTGTAAGAAGTAAAGTATTCAGTTATTCAGAAAGTA	6369
Qy	8146	ACAGCACTTCTTAAGCAGTTAAACAGGTTCTGTATTGGAGAGAGTTGAGATCCAAAGCA	8205
Db	6370	GTAAAGAGAAAGAGATCAGTTAATGAAAGATTGAAGAAAGATATCAGTTTCTGAGAA	6429
Qy	8206	GAAGCTACTGTCGACAGTTAAACCTATGCTGTAAGATTGAGAAAGAAATTTGTTACAGACA	8265
Db	6430	GTTATGAAACAGAAATTTGTTACTGAAGAAAGTGAAGAAAGAAAGATTCAGTTGTTGAAGA	6489
Qy	8266	AATGATTTGAATGCAATTTCTTAAAGTAGAAGTTTGATGATGATTAAGTAACTGCT---	8322
Db	6490	ATTGTAGAGGAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAGAGAGATCAGTTGTT	6549
Qy	8323	-----AATCTTCTGTATGATCAAGAAATCGAGAGAAATTCGCGAGCGAGGTAAAT	8377
Db	6550	GAAGAAATTTGAGAGGAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAGAAAGATCA	6609
Qy	8378	CTTGACAGCAACAAGTAACTGTAATCCGTAGTTCGTTTACCAAGCAAGATTAATGAAA	8437
Db	6610	GTTGTTGAGAAATTTGAGAGAGAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAGAA	6669
Qy	8438	ATTAATGATTAACAAAAAAATATATTTCA--GAATCAATGCTCTTGTTAAATGATACA	8496
Db	6670	GGATCAGTGTGTAAGAAATTTGAGAGAGAAAGATCAGTTGTTGAAGAAATTTGATAG	6729
Qy	8497	AAGATGAGCGAAATATAGAAATCTTTAGCGGTAGCCGTGTGATGCACAAGAACAAAC	8556
Db	6730	GAGGAAGGATCAGTTGTTGAAGAAATTTGAGAGAGAAAGATCAGTTGTTGAAGAAAT	6789
Qy	8557	AAAGATTTAGAGATCAAAACAGTTAATCTTCAACTGTAATGAGAGAAAGATATCT	8616
Db	6790	GTAGAGAGAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAGAAAGATCAGTTGTTGA	6849
Qy	8617	CAACTTCGTGCAAAAAGCTTTGGCTTAAATATGAAATTTATGAAATGTTAAAGAACTGGA	8676
Db	6850	GAAATTTGAGAGAGAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAGAAAGATCAGTT	6909
Qy	8677	GAGCCTTAGTCGAGCGGAAACGACGCCGTTGAAATTTATCAAAAGATCTACAGGA	8736
Db	6910	GTTGAAGAAATTTGAGAGAGAGAAAGATCAGTTGTTGAAGAAATTTGAGAGAGAAAGGA	6969
Qy	8737	GCATTCGTTGCAGAAATTTGGGAAATTTGAGATTAATTTAGAAACGATTGCAAGATTAAT	8796
Db	6970	TGAGTTGTTGAAGAAATTT--GTAGAGAGAGAGATCAGTTTGTGAAGTATGTTGATGAA	7026
Qy	8797	ACGATTTGTAAGATCAACGAGACGGAACCAAGAGGTCCTTGTGGAAGAAATGATATT	8856
Db	7027	ACTGAAATTTAGTAAATGATGAAATTTAGAACAACTCATTTACTGAAAGATTGAAGAA	7086
Qy	8857	TCTGTGAAAAATACAATTTCAAGGGGAAACAAAATCATCTTGAAGATTAAGCCAGAAAT	8916
Db	7087	CAAGTATCAGTTAAGATGAAATATATAGAAATGATCAGTTGCTGAAGCGTTGAAGAA	7146
Qy	8917	GTTTGAACCGGAAGCTTAATATGTAGATGCTTTGAATGAACCTTGATGATCTACAAGGA	8976
Db	7147	TCAGAAATCTAATTCAGAAATCAGTTTCCCAAGAGAGAGAAACAGAAAAGATTTGTTATT	7206
Qy	8977	AAAAAGTGTGCTATGTCGTAATTTGTTGTTGAAATGTTGATGTTAAATATGATTAAG	9036
Db	7207	GAAGAAATTTTGAAGAAATCTGAATCCGTTATGCGGAAAAATATTAACAAAGAAATCAGAT	7326
Qy	9037	AAAAATGTGAAGCCAAATTCGGAAGACATGCTATTGTGAATCTACTGGAACCAAGAA	9096
Db	7267	GAAGAAATTTTGAAGAAATCTGAATCCGTTATGCGGAAAAATATTAACAAAGAAATCAGAT	7326
Qy	9097	TATCAAGCAATTTACAAGAGCAAAATTAATATTTCTTGAAAAAGGAGACGTCAGCTGCA	9156
Db	7327	GCAGAGAAATTTTGAAGAAATCTTAAGTGAAGAGTTGTAAGTCTCAAGCAGATCA	7386

```
QY 9157 GCTGCAATATCGAATGTACACATTTCCAAATGAGATGATATTTAAAAATTTGGCAAGCAG 9216
Db 7387 ACTTCGAAAGAAATAGTATAGAGAGAGAGATCAGTTACTTAAAGATTGAGAAAAAGAA 7446
QY 9217 TATGACATCTTCAATTAATTAACCAAAAAATTAATATTAATTAATTAATTAATTAATTA 9276
Db 7447 TCAGTACTGAGAAATTTAGTATGAGATCAGTAATCGAAGAAATTAATTAATGAGAA 7506
QY 9277 AATGATTCGAATGTGAATGT---TCATGGGGTGGCTGAAGCAAGAGGTCCAGAGCCAA 9333
Db 7507 TCAGTACTGAGAAAGTTGTAGAAACAAAGGGGTTCAATTCCTCAAGAAATAGTAGAAGA 7566
QY 9334 GCGACAGTTAGTGTAAAGATCAAAATTAATAGAACTTAATTAATTAATTAATTAATTAAT 9393
Db 7567 GAATCAGCTACTGAAAGAAATTAATCGGGATGAAACTTAATGTTGAAAGAGTCTTGAAGAA 7626
QY 9394 ATTTAAACAGAGGAAACATCAATGATATATGCCGATATGATTAATTAATTAATTAAT 9453
Db 7627 GAAGGATCTGCTACTGAGAAATAGTATCAAGATGATCAGTACTTAATTAATTAATTAAT 7686
QY 9454 AAGACAAATTTCT 9465
Db 7687 AACAAGGATCT 7698

RESULT 32
AC129389_1/c
WPCOMMENT
Sequence split into 5 fragments LOCUS AC129389 Accession AC129389
Fragment Name Begin End
AC129389_0 1 110000
AC129389_1 100001 210000
AC129389_2 200001 310000
AC129389_3 300001 410000
AC129389_4 400001 425691
Continuation (2 of 5) of AC129389 from base 100001 (AC129389 Ratius norvegicus clone CH2

Query Match 1.1%; Score 105.2; DB 14; Length 110000;
Beet Local Similarity 42.1%; Pred. No. 4,4e-05;
Matches 591; Conservative 0; Mismatches 806; Indels 7; Gaps 3;

QY 7105 GTACAGAAAAAGTAAATGTAAACGCAAAAGAGAAAAATTAATTAAGCAACACACACA 7164
Db 16649 GAAAGAGAAAGAAAGAAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16550
QY 7165 AATGCAAGATTCGAGAGACAGCAATCGAGCCAAATGCTTGTGTAATTAATTTTGAACA 7224
Db 16589 GAAAGCGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16530
QY 7225 GCTGTAGAAGATAGAAAAATTTCTGAAGAAAGAAAGAAAGAAAGTTTAAAACTTTAGAC 7284
Db 16529 GAAAGCAGCAAGAAAGCAAGAGCAGAAAGCAGAAAGCAGAAAGCAGAAAGCAGAAAGAA 16470
QY 7285 GAAAGTTAACAAAGAACAGATTAATAAAGTTAAATGATGCTACGAAAAAAATCTTACAATCA 7344
Db 16469 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16410
QY 7345 GCAGGTATTTTCTACAGAAAGATCTTCTGTAAAGCGATAGAGAGATCTCAGGAGAA 7404
Db 16409 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16350
QY 7405 GGAATTAAGCATGTGTGAAGCTTCTGTATTAATTTGAAAAAATGTGATATTAACAACA 7464
Db 16349 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAG 16290
QY 7465 GAGGACAAGAAATTAATATCACTTCTACTGGTGTGGAACTGCAAGCTCTTGCTCCGCA 7524
Db 16289 AAGAGAAAGAAAGAAAGCAAGAAAGAA- GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16231
QY 7525 TCAGAACAGTGGCAGTTACAAATTAATTAAGAAATTTCCGAGATTACTGTTGAAATTTCT 7584
Db 16230 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 16171
```

```
QY 7585 TTTGTGAAAGCAGCTGAAAGAAAGTAAATGTAGATCGATATTTACAGAAATGTGCTTTA 7644
Db 16170 NNNNNNNGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCTTTTGAAGA 16111
QY 7645 ACAGCATATCAAGGTCTGTAGAGACATTTGGAAATAGAGAGTGTCTATGACGAATTTAAAT 7704
Db 16110 AGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 16051
QY 7705 TCTAATGGAAGATCAAAATATCATGATTAATTAATTTCTAAGCTATTAGAAAAAATTTGAT 7764
Db 16050 GAAAGAAAGAAAGCAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAGAAAGCAGAA 15991
QY 7765 GTTATTTGTTAAAGATTAATTCGGAATTGAGAGCGGAAAGCAAAAGCTTAACCTAGAGCG 7824
Db 15990 GCAGAAAGCAGAAAGCAGAAAGCAGAAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAA 15931
QY 7825 GTAGCTGCCGAGACCATATATCTCAAAAGCAAAAGAAAGAAATGAAATTCAGAGGTTGAAATT 7884
Db 15930 GAAAGAAAGAAAGAAAGAAAGATTTTGTAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAA 15871
QY 7885 GAGAAAGTATTTTCAATGAAGAAATAGAGTAACTAGCCCTTTAAAGCAATTTGAAAGA 7944
Db 15870 GCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 15811
QY 7945 GAATCAATGTCAAAAGTGAAGAAAGAAAGAAAGAAAGTGAATCTCAAGAGCTTCT 8004
Db 15810 GAAAGAAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 15751
QY 8005 GTAGAGCAGATGACAGGGGCGAGAAATTTTCCGAAAGCAAAAGATGCCGAGAGCTTAT 8064
Db 15750 GAAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAGAAAGAAAGAAAGAA- 15692
QY 8065 TTGAAAGTTAGTACAAAATCCGAAAGAAATTTTTCATGCAATATGTAATATGGA 8124
Db 15691 ---AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 15635
QY 8125 GCAACACTTAATTAATGAAGTAAACAGCAAGTTTCTTAAGCAGATTAACAGGTTCTGTATTGGA 8184
Db 15634 GAAAGAAAGAAAGCAGAAAGAAAGAAAGAAAGAAAGCAGAAAGCAGAAAGCAGAAAGAA 15575
QY 8185 GGAATGGAAGTCAACCAAGCAGAAAGCTACTGCTCAGAGTTAAATCTATGTGAAGTTAG 8244
Db 15574 GCAGAAAGAAAGAAAGAAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 15515
QY 8245 GAAAGAAATTTGTTCAAGAACTGAATGATGCAATTTCTTAAAGTGAAGGTTTGAT 8304
Db 15514 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 15455
QY 8305 GAAATTAAGTAACTGCTTAAATCTTCTGTAGTATCAGAAATGAGAGGAATTTGCCGA 8364
Db 15454 GAAAGAAAGAAAGCAGCAGAAAGAAAGCAGAAAGAAAGAAAGAAAGCAGAAAGAAAGAA 15395
QY 8365 GCAGAGTGAATTAATCTTCA--CAGACAAAGTAAATTAATCTGAATCCGTGTTTGAAGA 8422
Db 15394 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 15335
QY 8423 AGCAAGATTATGAATTAATGATTAACCAAAAAATATATTTCAAGAGTCAATGCTCTTG 8482
Db 15334 AAGAAAGAAAGAAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAGAAAG 15275
QY 8483 CTTTAAATGATCAAAAGATGAAG 8506
Db 15274 CAGAAAGAAAGAAAGAAAGAAAG 15251

RESULT 33
E10125
LOCUS E10125 2069 bp RNA linear PAT 29-SEP-1997
DEFINITION DNA encoding an immunogenicity protein.
ACCESSION E10125
VERSION E10125.1 GI:22026753
KEYWORDS UP 1995284392-A/1.
```

SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 2069)  
AUTHORS Doi,H., Nagakuchi,Y., Tanaka,Y. and Fujiieki,Y.  
TITLE GENE CLONE MANIFESTING CHICKEN LEUCOCYTOZON IMMUNOGENIC PROTEIN  
AND GENE RECOMBINANT VACCINE AGAINST CHICKEN LEUCOCYTOZON  
JOURNAL Patent: JP 1995284392-A 1 31-OCT-1995.  
COMMENT DOUBUTSUYOU SEIBUTSUGAKUTOKI SEIZAI KYOKAI, KITASATO INST:THE  
OS Leucocytozon caulleryi  
PN JP 1995284392-A/1  
PD 31-OCT-1995  
PF 19-APR-1994 JP 1994080643  
PI DOI HIROHITO, NAGAKUCHI YOSHIO, TANAKA YOSHIO, FUJISAKI YUUIRO  
PC C12N15/09,A61K39/015,C12P21/02;  
CC strandedness: Double;  
CC topology: Linear;  
FH key Location/Qualifiers  
FT source 1..2069  
FT /strain='shizuoka'  
FT /organism='Leucocytozon caulleryi' FT  
FT misc\_feature 1..25  
FT /note='EcORI adaptor'  
FT misc\_feature 2045..2069  
FT /note='EcORI adaptor'  
FT misc\_feature 25..2046  
FT /product='immunogenicity protein'.  
FEATURES  
source Location/Qualifiers  
1..2069  
/organism="unidentified"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 1.1%; Score 104.2; DB 6; Length 2069;  
Best Local Similarity 43.3%; Pred. No. 0.00026;  
Matches 748; Conservative 0; Mismatches 963; Indels 16; Gaps 5;  
QY 7110 AGAAAAAGTAAATGTTAAACGCAAAAAGGAAAAAATATTAAAGCAAAACAGCAGCAATGC 7169  
DB 312 AAACACACATCTGAACACGAAATTAACAATCTTAATGAAGAAAGATTAACAAATGA 371  
QY 7170 AGGAATCGAGAGCAGCAATCGAGCCAAATGCTTGTAATTAATTTTGAACAGCTGT 7229  
DB 372 TGAAGACGAAACAACATTAATGAAGAAATTAATTAATAGATGATTTTACAGA 431  
QY 7230 AGAAGATGAAAAAATTTCTGAAGAAAAAGAACAGAAAGTTTAAAACTTAAACAGAGT 7289  
DB 432 AGTAAATGTAGAAAGATCAGATGAAACAGAAACATTAATGAATGAATGAATGAATGAT 491  
QY 7290 TAAACAAGAACAGATTAATAATGATGCTACGAAAAAATCTTAACATCAGCAG 7349  
DB 492 ACAAGAAAGCATGAAAAATTAACATGAAGAAAGAAAGAAAGTAAACATGAAGA 551  
QY 7350 TATTTTACAGAGATATCTTCTGTAAAAAGCGAATAGAGATATCTCAGGAGAGAGAT 7409  
DB 552 AATAGAAAAAGAGAGATGAAGATTAATACATGAAGAAAGAAAGAAAGTAAACA 611  
QY 7410 TAAAGCATTTGTGAAGATTTCTGATATTATTGAAAAAATGTGATTTTAAACAGAGA 7469  
DB 612 TGAAGAAATGAAAAAGAAAGAGATGAAGAAATTAATCATGAAGAAAGAAAGAAAGT 671  
QY 7470 CAAGAATATATCATCTTCTAGTGGTGTGGGAATGACAGTCTTGCTCCGATCAG 7529  
DB 672 AACACATGAAGAAAAAGAAAGAGCATGAAGAAATTAATCATGAAGAAAGAAAGAA 751  
QY 7530 AACAGTGCAGTT--ACAAATATTAAAAAGAAATTCGGAAGTTACTGTTAAAAATCTTTT 7587  
DB 732 AGAAGTAAACATCAAGAAATTAAGAAAAAGAAAGCAAGAAAGTAAATACATGAAGAGA 791  
QY 7588 GTGAAGACACTGAAAAAGTAAATGTTAAGTCCGATTTACAGGAATGTTGCTTTACA 7647

DB 792 AAAAGAAAGTAAACATGAAGAAATAGAAAAAGAAAGCATGAAGAAATATACATGA 851  
QY 7648 GCATTCAGAGTCCCTGTAGAGCATTTGGGAATAGAGCTGCCTATGACGAATTAATTC 7707  
DB 852 GGAAGAAAAAGAAAGATTAACATGAAGAAATGAAGAAAGAGCATGAAGAAAT 911  
QY 7708 AATGGAATCAATCAATATCATGATTTAAAAATTTCTAAGTATTAGAAAAAATATTGAT 7767  
DB 912 ACATGAAGAAAGAAAGAAAGAAAGTAAACATGAAG--AAATGAAGAAAGAAAGCATGA 969  
QY 7768 ATTGTAAAAAGTAAATCGGAATTGAGCGGAGCAAAAGATTAACCGTAGAGCGGTA 7827  
DB 970 GAAGTAAATACATGAAGAAAGAAAGAAAGATTAACATGAAGAAAGAAAGAAAG 1029  
QY 7828 GCTGCCGAGCATTTATCTCAAAAGCAAAAGATGAATGAATTCAGAGTTGAATTTGAG 7887  
DB 1030 CATGAAGAAAGTAAATCATGAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAG 1089  
QY 7888 AAGATATTTTCAATGAAGAAATAGAGTAACTAGCCCTTTAAAGAAATTTGAAGAGA 7947  
DB 1090 AAGATTAACATGAAGAAAGAAAGAAAGAAAGATTAATCATGAAGAAAGAAAGAAAG 1149  
QY 7948 ATCAATGTCMAAGTGAAGAAAGAAAGAAAGAGTGAATCTCAAGAGCTTCTGTA 8007  
DB 1150 ATACATGAAGAAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGAAAGTAAACAT 1209  
QY 8008 GGAGCATGACAGGGCGAGAAATTTTCCGAAGCAAAAGATCCGGAAGCTCTTA---- 8063  
DB 1210 GAAGAAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGAAAGTAAACATGAAGAA 1269  
QY 8064 -TTTGAAGTTAGTCAAAATCCGGAAGATTTTTCATGCA-GATTAATGAAATATG 8121  
DB 1270 GAAAAAGAAAGTAAACATGAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGAA 1329  
QY 8179 ---TTGGAGAGTTGGAGTCCCAAGCAGAAAGCTAAGTAACTATGTA 8235  
DB 1390 CATGAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGAAAGTAAATGAAGAAAG 1449  
QY 8236 GAAGTTGAAGAGAAATTTGTTCAAGCAAAATCGATTGAATGCAATTTCTAAAGTAA 8295  
DB 1450 GAAAGAAAGAAAGAAAGAAATTAACATGAAGAAAGAAAGAAAGTAAACATGAAGAAAG 1509  
QY 8296 GGTGGATGAAGATTAAGTAACTGCTAAATCTTGTAGTATCAGAAATGAGAGAG 8355  
DB 1510 GAAAAAGTAAACATGAAGAAAGAAAGAAAGTAAATCATGAAGAAAGAAAGAAAG 1569  
QY 8356 ATTGCCGAGCAGAGTGAATCTTCTACAGCAAAAGTAAATCTGAATCCGTAATCGT 8415  
DB 1570 GATGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTAAAG 1629  
QY 8416 TTAAGAAACAGATTAATGAATTAATGATTAACAAAAAATTAATTTCAAGAGTCAAT 8475  
DB 1630 GAAGAGAAAGAAAGAAAGAAAGAAAGTAAAGAAAGAAAGAAAGAAAGTAAAGAA 1689  
QY 8476 GCTCTTGCTTTAAATGATTAAGAAATGAAGCAATATGAATCTTAAAGCGTAGCCGT 8535  
DB 1690 GAAAGAAAGAAAGAAAGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1749  
QY 8536 GTGCATGACAAAGAAACAAAGCAATTAAGAGATCAACAAAGTTAACTTTTACAACT 8595  
DB 1750 GTAAACATGAAGAAAGAAAGAAAGTAAACATGAAGAAAGAAAGAAAGTAAACAT 1809  
QY 8596 GTAAATGAGAGAAAGTATCTCACTTCGCAAAAGCTTTGCTAAATGAATTAAT 8655  
DB 1810 GAAAGAAAGAAAGAAATTAATTAAGAAAGAAAGAAAGAAAGTAAACATGAAGAAAG 1869  
QY 8656 GGAATGTAAGAAAGAACTGAGAGAGCTTAAGTGAAGCGGAAACAGACGCTTTGAAT 8715  
DB 1870 GAAAAAGTAAACATGAAGAAAGAAAGTAAACATGAAGAAAGAAAGTAAAGTAAAG 1929

OY	8716	TATTCAAAGACTGACCTACGCGAGCATTGGTGTCCAGAAATTCGAAATTCGAGATAAATTGA	8715
DB	1930	TATGAGAAGAGAGAGAAAAAGTAACATCAATGAGAGAGAGAGAAAAGTAATGAGAG	1989
OY	8776	GAAACGATTGCAGAGAGATATATCATTTGTAAGCTCAACGAGACGC	8822
DB	1990	AAGAAAGATTAATGAAAGTATATAAAAAAAAAAAAAAAAAAAAAAAAAAGG	2036
<hr/>			
RESULT 34			
LOCUS	E10126	3399 bp	DNA linear PAT 29-SEP-1997
DEFINITION	DNA encoding an immunogenicity protein of leucocytozon cailleryi		
ACCESSION	E10126		
VERSION	E10126.1	GI:22026754	
KEYWORDS	JP 1995284392-A/2.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 3399)		
AUTHORS	Doi H., Nagakuchi Y., Tanaka Y. and Fujisaki Y.		
TITLE	GENE CLONE MANIFESTING CHICKEN LEUCOCYTOZON IMMUNOGENIC PROTEIN AND GENE RECOMBINANT VACCINE AGAINST CHICKEN LEUCOCYTOZON		
JOURNAL	Patent: JP 1995284392-A 2 31-OCT-1995; DOUBTSUYOU SEIBUTSUGAKUTEKI SEIZAI KYOKAI, KITASATO INST:THE		
COMMENT	OS None OC Artificial sequences. PN JP 1995284392-A/2 PD 31-OCT-1995 PF 19-APR-1994 JP 1994080643 PI DOI HIROHITO, NAGAKUCHI YOSHIO, TANAKA YOSHIO, FUJISAKI YUJIRO PC C12N15/09,A61K39/015,C12P21/02; CC strandedness: Double; CC topology: Linear; FH Key Location/Qualifiers FT source 1..3399 /organism='Artificial sequences' FT CDS FT 1..3399 /product='fusion protein of maltose-binding protein and an immunogenicity protein' FT misc_feature 1..1149 /note='maltose-binding protein' FT FT misc_feature 1150..1174 /note='ECORI adaptor' FT misc_feature 1174..3195 /note='immunogenicity protein' FT FT misc_feature 3194..3218 /note='ECORI adaptor' FT misc_feature 3219..3399 /note='sequence derived from pMAL-C vector'. LOCATION/Qualifiers 1..3399 /organism='unidentified' mol_type='genomic DNA' db_xref='taxon:32644'		
<hr/>			
ORIGIN			
Query Match	1.1%	Score 104.2;	DB 6; Length 3399;
Best local similarity	43.3%	Pred. No. 0.00022;	
Matches	748;	Conservative 0;	Mismatches 96;
		Indels 16;	Gaps 5;
OY	7110	AGAAAAAGTAAATGTATACGCCAAAAAGAGAAAAAATTTAAGCAACAGCAAGCAATGC	7169
DB	1461	AAACAACAATCTGGAACAAGAAATTAACAATACCTAATGAAAAACGATATTAACAAATGA	1520
OY	7170	AGGAATCGAGAGAGAGCAATTCGAGCCAATGCTTGTTAAATTAATTTTGGAAACAGCTGT	7229
DB	1521	TGAAGAACCAACACACTTATGAAAGAGAAATTAATTTATATAGTAGATGATTTACAGCA	1580

QY	7230	AGAAAGATAGAAAAAATTTCTGAAGAAAAAGGAA	CGAAAGTTTAAAAA	CTTTAGACGAAGT	72829
Db	1581	AGTAAATGTGAGAAATTCAGTGAACAGAACATAT	TAACTATGAAATTAAGTAA	TGACAT	1640
QY	7290	TAACAAAGAACAGATTAATAAAGTAAATGATGCT	GCACAAAAAATCTTACATCAGCAGG		7349
Db	1641	ACAGAAAGAGCATGAAAAAGTAAACATGAAAGAA	AAAAAGAAAGTAAACATGAAAGA		1700
QY	7350	TATTTCTACAGAGATCTCTGTAAAAACGGATAG	GAGATCTCAGGAGAAAGAT		7409
Db	1701	AATGAAAAAAGAAAGCATGAAAGATTAATCA	TGAAAGAAAAAAGAAAGTAAACA		1760
QY	7410	TAAAGCAATGTGAAAGCTCTGATATTATTTG	AAAAAATGTGATTTCAACAGAGGA		7469
Db	1761	TGAAAGAAATGAAAAAGAAAGACATGAAAGAT	ATATCATGAAGAAAAAGAAAGAGT		1820
QY	7470	CAAGAAATATATCACTTCTACTGGTGGTTTG	GGAATGCAGTCTTGCCGATCAGG		7529
Db	1821	AACACATGAAGAAAAAAGAAAAAGAGCATG	AAAGATTAATCATGAAGAAAAAGAA		1880
QY	7530	AACAGTGGCAGTT--ACAAATATTTAAAGAA	ATTTCCGAGTTACTGTTGAAAAATCTTTT		7587
Db	1881	AGAAGTAAACATGAAGAAATATGAAAAAG	AGACATGAAGAGTAAATACATGAAGA		1940
QY	7588	GTGAAGACAGCTGAAAAAGTAAATGTTAGAT	CGGATTTACAGAAAAATGTTGCTTTAACA		7647
Db	1941	AAAAAGAAAGATTAACACATGAAAGAAAT	TGAAAAAAGAAAGACATGAAGATTAATACATGA		2000
QY	7648	GCATATCAAGGCTCGTAGAGCACTTGGGAA	TAGAGCTGCTATGCGAATTTAAATTC		7707
Db	2001	GGAAAGAAAAAGAAAGTAAACATCACTG	AAAGAAATGAAAAAGAAAGACATGAAGAGTAA		2060
QY	7708	AATGAAAGATCAAAATATCAGTATTAAAT	TTCTAAGCTATTAGGAAAAAATATTTGATGTT		7767
Db	2061	ACATGAAAGAAAGAAAAAGAAAGTAAAC	ACATGAAG--AAATAGAAAAAGAAAGACATGA		2118
QY	7768	ATTGTTAAAGATTAATTCGGAAATTGAG	AGCCGAGAACAAAGATTTAACCTTAGAGCCG		7827
Db	2119	GAAGTAAATCACTGAAAGAAAAAAGAA	AGTAAACACATGAAGAAAAAAGAAAGAG		2178
QY	7828	GCTGCCGAGCCATTATCTCAAAAGCAA	AGAAATGAATTCAGAGGTTGAAATTTGAG		7887
Db	2179	CATGAAGAGTAAATCACTGAAGAAAGAA	AGAAAGATTAACACATGAAGAAAGAAAGAA		2238
QY	7888	AAGAGTATTTTCAATGAAAGAAATAG	ATGATCTAGCCCTTTAAAGCAATTTGAAAGAA		7947
Db	2239	AAAGTTAACATGAAAGAAAGAAAGAA	AGAAAGTAAATACATGAAGAAAGAAAGAA		2298
QY	7948	ATCAATGTCAAAGTGGAAAAAGAAAC	AGAGTGACTGCTGAATTTCAAGAGCTTCTGTA		8007
Db	2299	ATACATGAAGAAAGAAAGAAAGATTA	CACTAGAAAGAAAAAGAAAGTAAACAT		2358
QY	8008	GGAGCAGTAGAGGGGACAGAAATTA	TTTCGAGACAAAGATGCGCGAGCTCTTA----		8063
Db	2359	GAAGAAAGAAAAAGAAAGATCA	CATGAAGAAAAAGAAAGATTAACATGAAGAA		2418
QY	8064	-TTTGAAGTTAGTACAAAATCCG	AGAAAGTATTTTTCATGCA--GATTAATGTGAATATG		8121
Db	2419	GAAAAAGAAAGATTAACATGAAGAA	AGAAAGAAAGTAAACATGAAAGAAAGAAAGAA		2478
QY	8122	GAAGCAACACATTAATGAAGAAAGT	ACAGCAAGCTTTCTAAAGCATCAAGTTCTGTA--		8178
Db	2479	GAAGTAAACACATGAAGAAAGAA	AGAAAGTAAACACATGAAGAAAGAAAGAAAGTAA		2538
QY	8179	---TTGGAGAGAGTTGAGATCA	CCAAAGCAGAAAGTACTGCTGCAAGTAAACATATGTA		8235
Db	2539	CATGAAGAAAGAAAGAAAGTAA	CACTGAAGAAAGAAAGAAAGTAAACATGAAGAAAGAA		2598
QY	8236	GAAGTTGAGAGAAAGAAATTTG	TTTCAAGAAACAAATCGATTTGCAATTTCTTAAAGTAA		8295
Db	2599	GAAAGAAAGAAAGAAAGTAA	CACTGAAGAAAGAAAGAAAGTAAACATGAAAGAAAGAA		2658
QY	8296	GGTTTGATGAATAAAGTAACTG	CTAAATCTTCTGTAGTATCAGAAAAATGAGAGAGGA		8355

[illegible]

**AUTHORS** Wilson.R.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (13-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Oct 11, 2002 this sequence version replaced g1:22476311.

**COMMENT**

-----  
Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: submissions@watscn.wustl.edu  
Summary Statistics  
Center project name: M\_BB0466A03  
-----

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

**MAPPING INFORMATION:**  
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu

**SOURCE INFORMATION:**  
The RPCI-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

**NEIGHBORING SEQUENCE INFORMATION:**  
This sequence is the entire insert of the clone.  
Location/Qualifiers

**FEATURES**

**source**

1..179685  
    /organism="Mus musculus"  
    /mol\_type="genomic DNA"  
    /db\_xref="taxon:10090"  
    /chromosome="14"  
    /map="14"  
    /clone="RP24-465A3"  
    /clone\_1id="RPCI-24"  
    1392..1452  
        /rpc\_family="B4"  
    1489..1616  
        /rpc\_family="ERV.L"  
    2081..2289  
        /rpc\_family="ERV.K"  
    2311..3896  
        /note="Unresolved simple sequence repeat."  
    3884..4165  
        /rpc\_family="ERV.K"  
    3919..4254  
        /note="Sequence derived from one plasmid subclone."  
    4752..4976  
        /rpc\_family="L2"  
    7171..7369  
        /rpc\_family="MER2\_type"  
    7418..7585  
        /rpc\_family="MER2\_type"  
    7734..7942  
        /rpc\_family="L1"  
    7981..8384  
        /rpc\_family="L1"  
    8584..8891



[illegible]

RESULT 36					
LOCUS	AC133034				
DEFINITION	AC133034	233713 bp	DNA	linear	HTG 19-NOV-2002
ACCESSION	AC133034				
VERSION	AC133034.2	GI:25073581			
KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
	Eutherota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 233713)				
AUTHORS	Muzny D, Merle, Metzker, M Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alpebrooks, S., Amin, A., Angiano, D., Anyalelechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Behanmed, F., Bielwalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Butrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesaar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Ande, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Giddy, M., Guerra, W., Guevara, M., Guneratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harrey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladik, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huayk, S., Hune, Y., Idlebird, D., Jackson, A.,				

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karchh, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowib, C., Kraft, C.L., Ledow, H., Levam, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorensehu, L., Loueseged, H., Lozad, R., J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindartine, M., Mahmood, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
 Mawhney, S., McLeod, M.P., McNeill, T.Z., Meonen, E.,  
 Mlloasvillavic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
 Nankervils, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Naackeleh, O., Okunou, G., Olarunpasegon, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C.,  
 Piopier, F., Poindexter, A., Popovic, D., Primm, E., Pu, L., L.,  
 Piaz, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
 Riley, B., Reilly, M., Ren, Y., Roemer, M., Richards, S., Riggs, F.,  
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,  
 Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C., Smales, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
 Wang, O., Wang, V., Villaaana, D., Waldron, J., Walker, B., Wang, D.,  
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstein, G., and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 233713)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (06-SEP-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 23713)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department

\* NOTE: Estimated insert size may differ from sequence length



Web site: <http://www.sanger.ac.uk>  
Contact: [zfish-help@sanger.ac.uk](mailto:zfish-help@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Chitong Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see [http://www.sanger.ac.uk/Projects/D\\_zerio/fishmask.shtml](http://www.sanger.ac.uk/Projects/D_zerio/fishmask.shtml) DKRY-56E5 is from a Zebrafish BAC library

VECTOR: pindigBAC-5.

```
Location/Qualifiers
1. 228713
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-56E5"
/clone_1ib="Daniokey"
```

## ORIGIN

[illegible]

QY	7530	TAACGTGCGACGTTCAAAATATTTAAAGAAATCCGGAAGTTCTGTGTGAAATTTCTTTGT	7589
Db	152437	AAGA-----AGATTAAGTTGATGAGAGAGATGATCGAAGAAAGAAATTAAGTGAAGATGA	152383
QY	7590	GAAAGCAGCTGMAAAAGTAAATGTTAGTCGATATTAACAGAAATGTTCCTTTAAACGC	7649
Db	152382	AAAGTTTATGAAGATGAAGATGCTGAGAGAGA--AGAAAAAAACATGAAATCATGTAGT	152324
QY	7650	ATATCAAGTCCGTGAGAGCATTTGGAAATAGAGCTCCATATGAGAAATTAATTTCTAA	7709
Db	152323	GGATCCAGAAATAGAAAGATGATGCGAAAGAGGACAGATGAAGTATGAGAGATGATGA	152264
QY	7710	TGGAGATCAAAATATCAGATTTAAAAATTTCTTAAGCTATTTAGGAAAAAATATTGATGTTAT	7769
Db	152263	TGAAAAAGAACATTTGATGAAAGAAAGATTTGATCGAAGAAAGAAATTAAGAGAAAGATTA	152204
QY	7770	TGTAAAGATAAATCGGAATTGAGACGGAGCAAAAGATTAACCGTAGAGCCGTAGC	7829
Db	152203	GGTTTATGAAGAAATCAAGTTAAAGAAAGATCAGAAAGAGTAAAAAATGAAAGAGTTCA	152144
QY	7830	TGCCGAGCCATTATCTCABAAAAGAAAGATGAATGAAATGAAATTCAGAGGTTGAAATTGAGAA	7889
Db	152143	TGAACACGATTAAGAACAGATGCGAAGATTAAGATTAATGCGAG--AAAGTGCAGATGA	152085
QY	7890	GAGTATTTTCATGAGAAAGAAATAGAGTAACTAGCCCTTTCTAAAGAAATTTGAAAGAAAT	7949
Db	152084	AGGTGATGATATGATGAAAGAAATTTGATGMAAGACAGATGATGAGATGAGAAAGTTAA	152025
QY	7950	CAATGTCAAAGTGAAAAAGAAACAGAGTGACTGTGAATTTCAAAGAGCTTCTGTAG	8009
Db	152024	CAAGAGTATGTAGTATGATGCAAAAA---AGTACCAACGAAAGAAAGATGATGAAAA	151969
QY	8010	AGCAGTAGCAGGGCAGGAATTATTTCCGAGCAAAAGATGCCGGAAGCTCTTATTTGAA	8069
Db	151968	AGAACAGATGAGAGAACAGAGATGATGAAGAAAGAAATATGATGCTGAAGAAAGAAATATG	151909
QY	8070	AGTTAGTCAAAATCCGGAAGATATTTTTCATGCAATTAATGCAATATGGAACCAAC	8129
Db	151908	AAGTGAAGAAATGATGTCAGAACAGTAAATTAAGAGAGAGATGAAGAAAGTTGATTAAGAAAG	151849
QY	8130	ACATTAATAAGAAAGTAAACAGCAAGCTTTCTAAACAGTAACAGGTTCTGATTTGGAGAGAT	8189
Db	151848	AAATTTAGAT--CATGCGAAGAAAGAAAGAAAGATGAAGATTTGAAGATGATGAAGAGAA	151790
QY	8190	TGGAATCACCAGAGCGAAGTACTGCTCAGGTAAACCTATGCTGAAGTTGAGAGAG	8249
Db	151789	TGATGACGAAAGAAAGAAAGATGMAAGATTTGATGAAGAGAAAGAAATTTGAAGATGAAG	151730
QY	8250	AAATTTGTTCAAGAACAAATCGATTGATGCAATTTCTAAAGTGAAGGTTTGGATGAAG	8309
Db	151729	AAATGATGAAGAAAGATGTAGATTAATGAAGTTAAGT--TGATCCGAAAAAGAAATTAAGAG	151671
QY	8310	TAAAGTAATCGCTAAATCTTCTGTGATTCAGGAAATGAGAGAGAAATTTGCCGAGCAG	8369
Db	151670	AGATGAGATTGATGAAGAAAGATGAACCTTTAAAGAGATGCAAGAAAGATTAAGAAATGAAG	151611
QY	8370	AGTGAATCTTCTTAACAGCAAAAGTAAATCTGAATCCGTAGTTCTTTTACGAAAGCAAG	8429
Db	151610	CGTTGATGAAGACGATTAAGAAAGAAAGATCAGAAATTAAGCAAGAAAGAAAGATAGC	151551
QY	8430	TTATGAAAAATATGATTTACAAAAAATATTAATTTAGAGATCAATGCTCTTGCTTTAA	8489
Db	151550	AGATGAAAGGTGATGAATGAAGAAATTTGATGAAGAGCAGATGATGAAGATGTTAA	151491
QY	8490	TGATACAAAGATGAAGCAATTAAGAAATCTTTAGCCGTAGCCGCTGTGCATGACAAAG	8549
Db	151490	CGTTTATGAAGGTGATGTGATGATGAAGATTAAGTTGATCAAGAAAAACAAAGATGAAGT	151431
QY	8550	AACAAACAAAGCTTTATCGAGATCAAAACAGTTTAATCTTTCTACAACTGTAAATGAGAGAA	8609
Db	151430	TGAAGTATGAAGAGATGATGCAAGAAAGATTAAGCAATGAAGAGATGCAAGAAATGAG	151371

```
QY 8610 CCTATCTCACTCTTGTGCAAAAAGCTTTGGCTTAAAAATGAAAATGTAATAAGG 8669
DB 151370 ACATGATGCAAGAAAGAGAGAGATGAAGAAATATGATGAAGCTTGTGAATTAAGAGA 151311
QY 8670 AACTGAGAGAGCCTTAGTGTGGAGCGGAAAACAGCAGCCGTTGAAAATTAACAAGAGTAC 8729
DB 151310 TGAATGAGAAAAGAAAAGATGAAGATGAAGAAAGTTGAAGAAAGATGAAGAGTGA- 151252
QY 8730 TACAGAGCATTGGTTGCAAGAAATTGGGAAAATTGGAGATTAATTGAAAACGATTGCAAG 8789
DB 151251 TGCAGAAAAAGAAACAGAGATGTTGTTGATGAACGATGAAAAAATTAAGAGAAATGAGAG 151192
QY 8790 AGATAATACGATTGTAGAGTCAACGAGACCGAACCAAGAGAGTCTTGTCCG- AAAGA 8848
DB 151191 ATGATGCAAAAAATAGTGCAGAGAGAGAAAGAAAGATGATGATGATGATGAAAAAGA 151132
QY 8849 ATGGTATTTCTGTGAAAATTAACAATTTCAGGGGAAAACAAATTCATCCATTGAAGATTAAG 8908
DB 151131 ACAAGATGATGAAGAAAGATATGATGCTGAAAAGAGAAAGATTAAGAAATGAAGAGTTG 151072
QY 8909 CCAGATTGTTGGAAACCGAGAGTGTAAATGTAGATCTTTGAATGAACCTTGATGTAGATC 8968
DB 151071 ATGAAAGAGATGTAGATGAAAAACAATTAAGTTGATTAAGAAAAAGAAAGAGAGAGAT 151012
QY 8969 TACAAAGAAAAGTGTGCTATGTGTGAATTGTGTAATGGAATGTTGAATGTAATG 9028
DB 151011 ATGATGAGAAAAGGAATGAAGAT- GAAGATGATGTGAAAAAAATTAAGAAAGATGAAG 150953
QY 9029 TGATTAAGAAAATGTAGAAAGCCAAATTCGGAAGACATCTTATGTAGAACTACTGGA 9088
DB 150952 TGGATGATTAAGAAAGTGAAGCTTTAAAAAAGATGACGAAAAGAGTAAAGATGAAGAGTTG 150893
QY 9089 AACCAAGATATCAAGCATTTAACAAGACGAAAAGTAAATTTCTTGAAAAGAGAGCGCTG 9148
DB 150892 AATAAGACATGAAGAAACAAGATGCAAGAAAAATTAAGAAATGAAGAGATTAAGTGA 150833
QY 9149 CAGCTGCAGCTGCATATGAAATGTACACATTTCCATGATGATTAATTAATTTG 9208
DB 150832 TAGGTGGGATGAGAAAGCTGATGAAGATTAATGACAAAGAAAGATCTAATAAGAAAAAG 150773
QY 9209 CAAAGCAGATGATCTTCTCAATTAATTAACCAAAATTCAAAATAATA 9259
DB 150772 AAAAATTAAGAAAGATGATGATTAATGATGCTGTAAGAAAGAGAAAGA 150722

RESULT 38
AF270648 4102 bp DNA linear INV 12-JUN-2000
DEFINITION Plasmodium falciiparum mature parasite-infected erythrocyte surface
antigen (MESA) gene, complete cds.
ACCESSION AF270648
VERSION AF270648.1 GI:8468620
KEYWORDS
SOURCE Plasmodium falciiparum (malaria parasite P. falciiparum)
ORGANISM Plasmodium falciiparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 4102)
AUTHORS Shan, Z. X., Yu, X. B., Fang, J. M., Li, X. R., Ma, C. L. and Wu, Z. D.
TITLE Cloning and expression of the mature parasite-infected erythrocyte
surface antigen gene of Plasmodium falciiparum isolate FCCI/HN
Unpublished
2 (bases 1 to 4102)
JOURNAL Direct Submission
AUTHORS Shan, Z. X.
TITLE Submitted (20-MAY-2000) Department of Parasitology, Sun Yat-Sen
University of Medical Sciences, 74 Zhongshan Road II, Guangzhou,
Guangdong 510089, P. R. China
JOURNAL Location/Qualifiers
FEATURES
Source
1..4102
/organism="Plasmodium falciiparum"
/mol_type="genomic DNA"
/isolate="FCCI/HN"
/db_xref="taxon:5833"
```

```
/country="China"
<1..>4102
/gene="MESA"
join(<1, 150,281..>4102)
/gene="MESA"
/product="mature parasite-infected erythrocyte surface
antigen"
join(1..150,281..4102)
/gene="MESA"
/codon_start=1
/product="mature parasite-infected erythrocyte surface
antigen"
/protein_id="AAF75554.1"
/translation="MEVLCRNLCYDKNNMMENEGNKVKYVNNSLKVMKFCCTI
ICVFILDIYTNCEPTVSGIKNNNDYVYVILSTETPEMSLEIMRFDREHLSIR
NYIECLRAPIIIDPLMGSVYTDKRNKCLQIKLLEMAQSEERRRQGSBNKQDIBI
RKKEKEYLMKLEBMDSDVEKAFRELQFIKLDRTRPRKRVNVGSKETDSKTD
ESKETDSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSK
TGESKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETGSKETRIY
BETKYTKITSEPRETENVKITRESKDRGNVSGPYKNSNNTVTSSESETKCLAKS
ENEGKLGEMNDGASNSBDPKLPBENENGTRESSETDQKREBNKADNKKGS
KKKKSFFQMLGCNFKCNKITDDEESTLVVKDQAKKHFRLERANTKENDKDK
LLGEGDKEDVREKNDQKXVLAGSDKDVREKNDQKXVLAGSDKDVREKNDK
DKVIGSEKTOKEIKERKVKYKCKKKVKKGIKENDTEGNDKVGPEIILBEVERI
KKOVEDGIXENDTEGNDKVGPEIILBEVERIKKOVEBGIKENDTEGNDKVGPEI
TBEVKEEIKKOVEBGIKENDTESKDKLIGOEIILBEVERIKKOVEBGIKENDTE
TBEVKEEIKKOVEBGIKENDTESKDKLIGOEIILBEVERIKKOVEBGIKENDTESK
KVIGOEIILBEVERIKKOVEBGIKENDTESKDKLIGOEIILBEVERIKKOVEB
QEKVEKVEKVEKVEKVKVKNRNNENKDNVIGELKEDNEKDTANKDVIQEK
KEKEVEKVEKDTANKDVIQOEVIILBEVERIKKVKVKNRNNENKDNVIGELME
DVNEKDTANKDVIQOEVIILBEVERIKKVKVKNRNNENKDNVIGELME
EVENKDTANKDVIQOEVIILBEVERIKKVKVKNRNNENKDNVIGELME
ESESSESESESESESESESESESESESESESESESESESESESESESESESESE
ESESSESESESESESESESESESESESESESESESESESESESESESESESESE
ESESSESESESESESESESESESESESESESESESESESESESESESESESESE
SKENHTLNTFFPMALNEDFRREFPHILSIHEDDTLMELKRLILYNFLFLYNPMNRK
AELDKFSSENVVHQLINVERIRMYENARGLNTVLDPILITFNVI FGDITMFKF
IDE"

ORIGIN
Query Match 1.0%; Score 100.4; DB 2; Length 4102;
Best Local Similarity 43.9%; Pred. No. 0.0064;
Matches 572; Conservative 0; Mismatches 721; Indels 9; Gaps 3;

QY 7105 GTACGAGAAAAAGTAAATGTAAACGGCAAAAGCAAAAAAATATTAGCAACACAGCA 7164
DB 2459 GTAAAAAAGTAAATTAAGAAAAAGTAAAGAAATTAAGAAATTAAGAAATTAAGT 2518
QY 7165 AATGCAAGAAATCGAGAGACAGCAATCGAGCAATGCTTGGTAAATTAATTTTGAACA 7224
DB 2519 AAAAGTAAAGTATGAGCAAGAAATTAATGTTGAAGTAAAGAAAGAAATTTGAAAA 2578
QY 7225 GCTGTAGAAAGTAAAAAATTTCTGAAGAAAAAGCAAGAAATTTTAAAACTTTAGC 7284
DB 2579 CAAGTGAAGAAAGGAAATTAAGAAATATGATCTGAAATGAAGTAAAGTGAAGACAA 2638
QY 7285 GAAGTTAACAAAGAACAAAGTAAAAAAGTAAATGATGCTACGAAAAAATCTTACAATCA 7344
DB 2639 GAAGTGAAGAAAGGAGATTTATAGAAAGGTCCGAAAAACAAAGATTAAGTACAAAA 2698
QY 7345 GCAGGATTTTCTACAGAGATCTTCT--GTAAAAAGCGATAGAGAGATCTCAGGGA 7401
DB 2699 CAGGAAAAAGTAAAGAAAGTTTAAAAAAGTAAAGAAAGTAAAGAAAGTAAAGTAA 2758
QY 7402 GAAGGAATTAAGCCATTGTGAAGCTTGTGATTAATTTGAAAAAATGTAGATTTTACA 7461
DB 2759 AAAAGAAATTAATTAAGAAAGAAAGAAAGATATGTGATGAAAAAGAAATATGAAGA 2818
QY 7462 ACAGAGCAAGAAATTAATACCTTCTGCTGGTGGTGGGAAACGCGAGCTTGCTCC 7521
DB 2819 GATGTTAATGAAGAAAGATCCGAAACAAAGATTAAGTATGAACAAAGAAAAAGAAAA 2878
QY 7522 GCATCAGAGACAGTGGCAGTTTAAATTAATTAAGAAATTCGAGATTAATCTGTTGAAAT 7581
```



Db 950 AATTAAATTGAAAAGTAGATGAAGAACTGCGAAGAAATTAATTGAAAAAGTAGATGAAG 1009  
Qy 7949 TCAATGTCAAATGCGAAAAAGAAAAGAGAGTACTGCTGAATCTCAAGAGCTTCTGTAG 8008  
Db 1010 AAGTGTGCGAAGAAATTAATTGAAAAAGTAGATGAATTAAGTAAAAAGTAGCTGAAG 1069  
Qy 8009 GAGCAGTAGCAGGCGCAGGAATTAATTCGGAAGCAAAAGATGCCGAGCTCTTAATTGA 8068  
Db 1070 AATTAGTGAAGAAAAGTAGATGAACAAGTAGCTGAAGAAATTAAGTGAAGAAAAGTAGATGAAC 1129  
Qy 8069 AAGTAGTACAAAATCCGGAAGAAAGATTTTTCATGCGAGATATGTGAATATGGAAGCA 8128  
Db 1130 AAGTACTGAAAGAAATTAAGTGAAGAAAGTAGAACAAGTAGTGAAGAAAGTAGCTGAAG 1189  
Qy 8129 CACATTAATTAAGAAATGACAGCAAGTTTCTAAAGCAGTAACAGTCTGTATTTGGAGAG 8188  
Db 1190 AAGTAGCTGAAGAAAGTAGATGAAGAAAGTAGAAGAAAGTAGAAGAAAGTAGCTGAAG 1249  
Qy 8189 TTGAGATCACCAAGCAGAGAGCTACTGCTGACAGTAAAACTATGTAGAGTTGAGAGAG 8248  
Db 1250 TAGCTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAGAAATTAAGTGAAGAAAG 1309  
Qy 8249 GAAATTTGTTCAGAAC--AAATCGATTGAATGCAATTTCTAAAGTAGAAGGTTTGATG 8305  
Db 1310 TAGATGAAGAAAGTAGCTGAAGAAAGTAGTGAAGAAAGAAAGTAGAAGAAAGTAGCTGAAG 1369  
Qy 8306 AAGATTAAGTAATCTCTAATCTTCTGTAGTATCAGAAATGAGAGAGAAATTTGCCGAG 8365  
Db 1370 TAGTTGAAGAAATTAAGTGAAGAAAGTAGCTGAAGAAAGTAGTGAAGAAAGTAGCTGAAG 1429  
Qy 8366 CAGGAGTAGAATCTTCTACAGCAACAAGTAATCTGAATCCGTAGTTCGTTTACAGAAAGC 8425  
Db 1430 TACTTGAAGAAATTAAGTGAAGAAAGTAGTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAG 1489  
Qy 8426 AAGATTATGAAGAAATTAAGTGAAGAAAGTAATTAATTTACAGAG 8470  
Db 1490 TAGTTGAAGAAAGTAGCTGAAGAAAGTAGTGAAGAAAGTAGTGAAG 1534

## RESULT 40

CS082924 LOCUS CS082924 1701 bp DNA linear PAT 18-MAY-2005  
DEFINITION Sequence 11 from Patent WO2005040203.  
ACCESSION CS082924  
VERSION CS082924.1 GI:66349524  
KEYWORDS  
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)  
ORGANISM Plasmodium falciparum  
Plasmodium falciparum  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
REFERENCE 1  
AUTHORS Drullh, P., Singh, S., see See, D. and Mejia, P.  
TITLE Msp-3-like family of genes  
JOURNAL Patent: WO 2005040203-A 11 06-MAY-2005;  
Infect Dis J (FR)  
FEATURES  
source  
1. 1701  
/organism="Plasmodium falciparum"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:5833"  
1. 1701  
/note="MSP 3.6"

## ORIGIN

Query Match 1.0%; Score 98.6; DB 6; Length 1701;

Best Local Similarity 42.8%; Pred. No. 0.0015; Mismatches 809; Indels 6; Gaps 2;

Matches 610; Conservative 0;

Qy 7052 CTATAAATAGTTCTGTAGTAGCTCGAGTTCAACATTAATCTGATTTGACTTCGTAAGAG 7111  
Db 110 CTTTAAAAAGAAATTAAGAAATGAATTTTAAATTAATCTTAATTTTAAAGAAAGCA 169  
Qy 7112 AAAAAGTAATGTAAACGCAAAAAGGAAAAAATTAATTAAGCAACAGCAAGCAATGCGAG 7171

Db 170 ATAAATTTAATATAGAAAAATGAATAATTAATACAAATTAATATGAAGTAAACAGAGAAA 229  
Qy 7112 GAATCGAGAGCAGCAATCGAGCAATGCTTGTGTAATTAATTTGGAAACAGCTGAG 7231  
Db 230 CTAAAGAAAGAAATTAATATATTAATTAATTAATTTTCCGTATTAATTTTCTTGATA 289  
Qy 7232 AAGATGAAAAAATTTCTGAGAAAAAGAACAGAAATTTTAAAACTTTAGACAGAGTTA 7291  
Db 290 TCTTTACTGAAAAATTAAGAACAAAAAATGAAGAGTACCAATGAAGAAATTAAGAAAGTAA 349  
Qy 7232 ACAAGAAACAGATTAATAAGTAATTAATGATCTGACAAAAAATTTTACATCAGACGTA 7351  
Db 350 ATGATGAGAAAGAAAGTAATAAAGCAATATGATCTGAAATAAATGAAGAAAGTGAAGAAATA 409  
Qy 7352 TTTCTACAGAAAGATCTCTGTAAAGGGAATGAGAGATCTCAGAGGAGAGAGAAATTA 7411  
Db 410 AATCGAAACTGAATTAAGTGAAGAAATTAATCTGAAGAAAGTAGATGAAGAAAGTGAAG 469  
Qy 7412 AAGCCATTGTGAGACCTTCTGATTAATTAATGAAAAATGATAGATTAACAACAGAGACA 7471  
Db 470 AAGTAGCTGAAGAAATTAAGTGAAGAAAGTAGAAGAAAGTAGAAGAAAGTAGCTGAAG 529  
Qy 7472 AGAATATATCACTTCTAAGTGTGTTGGCACTGCAGGCTTCTCCGATCAGGAA 7531  
Db 530 AAGTAGATGAAGAAAGTAGCTGAAGAAAGTAGATCAAAAAAGTAGATGAAGAAAGTGAAG 589  
Qy 7532 CAGTGGCAGTTACAAATTTTAAAGAAATTCGAGGATTAAGTTGAAATTTCTTTGGA 7591  
Db 590 AATTAAATTAAGAAAGTAGATGAAGAAAGTAGAAGAAATTAATTAAGAAAGTAGATGAAG 649  
Qy 7592 AAGCAGCTGAAGAAAGTAATGTTAGATCGATATATTACAGAAATGTTCTTTTACAGCAT 7651  
Db 650 AAGTTGCTGAAGAAATTAATTAAGAAAGTAGAAGAAAGTAGAAGAAAGTAGAAGAAATTAAG 709  
Qy 7652 ATCAAGTCTCTG--TAGAGCATTTGGAAATGAGAGCTGCTTATGACAGATTAATTTCTA 7708  
Db 710 AAGTACTGAGAAATTAATTAAGAAAGTAGAAGAAAGTAGAAGAAATTAATTTGAAA 769  
Qy 7709 ATGGAAGATCAAAATATCAATTAATTAAGAAATTTCAAGCTATTTGGAAGAAATTTGATGTTA 7768  
Db 770 AAGTAGCTGAAGAAATTAAGTGAAGAAAGTAGCTGAAGAAATTTGTTGAAAGTAGATGAAG 829  
Qy 7769 TTTGAAAAAGATTAATTCGAAATTTGAGAGCGGAAAGCAAAAGATTAACCTGAGAGCGTAG 7828  
Db 830 AAGTAGCTGAAGAAATTAAGTGAAGAAAGTAGAAGAAAGTAGAAGAAAGTAGAAGAAATTAAG 889  
Qy 7829 CTGCGGAGCAATTAATCTCAAAAGCAAAAGTAATGAATTAATTCAGAGGTTGAAATTTGAA 7888  
Db 890 AAGTAGATGAAGAAAGTAATCTGAAGAAATTAATTTGAAAAAGTAGATGAAGAAAGTAACTGAAG 949  
Qy 7889 AAGATATTTTCAATGAAGAAATTAAGTAATTAAGTAACTAGCCCTTTTAAAGAAATTTGGAAGAAA 7948  
Db 950 AATTAAATTAAGAAAGTAGATGAAGAAAGTAGCTGAAGAAATTAATTTGAAAAAGTAGATGAAG 1009  
Qy 7949 TCAATGTCAAATGCGAAAAAGAAAAGCAGAGTACTGCTGAATCTCAAGAGAGCTTCTGTAG 8008  
Db 1010 AAGTTGCTGAAGAAATTAATTTGAAAAAGTAGATGAATTAAGTGAAGAAAGTAGCTGAAG 1069  
Qy 8009 GAGCAGTAGCAGGCGCAGGAATTAATTCGGAAGCAAAAGATGCCGAGCTCTTAATTGA 8068  
Db 1070 AATTAGTGAAGAAAAGTAGATGAACAAGTAGCTGAAGAAATTAAGTGAAGAAAAGTAGATGAAC 1129  
Qy 8069 AAGTTAGTACAAAATCCGGAAGAAAGTATTTTCAATGCAATTAATGTAATTAAGAAAGCA 8128  
Db 1130 AAGTAGCTGAAGAAATTAAGTGAAGAAAGTAGAACAAGTAGTGAAGAAAGTAGCTGAAG 1189  
Qy 8129 CACATTAATTAAGAAATGACAGCAAGTTTCTAAAGCAGTAACAGTCTGTATTTGGAGAGAG 8188  
Db 1190 AAGTAGCTGAAGAAAGTAGATGAAGAAAGTAGAAGAAAGTAGAAGAAAGTAGCTGAAG 1249  
Qy 8189 TTGAGATCACCAAGCAGAGAGCTACTGCTGACAGTAACTATGTAGAGTTGAGAGAG 8248  
Db 1250 TAGCTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAGAAAGTAGCTGAAGAAATTAAGTGAAGAAAG 1309

```

QY 8249 GAAATTGTCAGAC--AAATCGATTGAATGCAATTTCTAAGTAGAAGGTTTGATG 8305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1310 TAGATGAAGAAAGTAGCTGAAAAAGTAGTTGAAGAAGAAAGTAGAAGAG 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8306 AAGATAAAGTAAGTCTTAATCTTCTGTAGTATCAGAAATGAGAGGAATTGCCGAG 8365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1370 TAGTTGAAGAAAGTAGATGAAGAAGTAGCTGAAAAAGTAGTTGAAGAAGAGTGAAAAAG 1429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8366 CAGAGTGAAATCTTCTACAGACAAAGTAATCTGAATCCGTAGTTCGTTACGAAAGC 8425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1430 TACTTGAAGAAGTAATTTGAAGAAGTAGTTGAAGAAGTAGCCGAAAGTAGCTGAAAAAG 1489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8426 AAGATTATGAAATATATGATTACACAAAAAATATATTTCAGAG 8470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1490 TAGTTGAAGAACAGGTGAAAAAGTAACAAAAATGATTTAATG 1534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 25, 2005, 14:17:05  
 Job time : 31065 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2005, 22:13:25 ; Search time 3407 Seconds  
(without alignments)  
19025.771 Million cell updates/sec

Title: US-10-647-057-8

Perfect score: 9726

Sequence: 1 acgagcggcaccatacaaa.....gagaaaaaggaaaaaatga 9726

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_21.\*  
1: genebegn1980s:\*  
2: genebegn1990s:\*  
3: genebegn2000s:\*  
4: genebegn2001s:\*  
5: genebegn2002s:\*  
6: genebegn2003s:\*  
7: genebegn2004s:\*  
8: genebegn2005s:\*  
9: genebegn2006s:\*  
10: genebegn2007s:\*  
11: genebegn2008s:\*  
12: genebegn2009s:\*  
13: genebegn2010s:\*  
14: genebegn2011s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9726	100.0	9726	6	AA167640
2	9718	99.9	11130	6	AA167647
3	2780	28.6	2780	6	AA167642
4	2322	23.9	2322	6	AA167645
5	2141	22.0	2141	6	AA167643
6	1887	19.4	1887	6	AA167644
7	1130	11.6	1130	6	AA167641
8	116.2	1.2	5361	6	AA178868
9	116.2	1.2	5529	6	ABK50883
10	116.2	1.2	6152	2	AA178867
11	104.2	1.1	3399	2	AA178866
12	98.6	1.0	1701	14	AD272262
13	87.4	0.9	3579	3	AA170099
14	81.8	0.8	1686	2	AAQ87587
15	80.4	0.8	1300	12	ADP85917
16	79.2	0.8	4965	8	ADA89806
17	79.2	0.8	7434	4	AA552179
18	79.2	0.8	7434	8	ABT14965
19	79.2	0.8	7434	8	ACF73459

20	79.2	0.8	7437	4	AA55232
21	77	0.8	204803	12	ADQ97348
22	76.4	0.8	4677	3	AA170259
23	76	0.8	1039	14	ACL64817
24	74.8	0.8	7104	4	AA551998
25	74.8	0.8	7107	4	AA554654
26	73.6	0.8	8155	2	AA174374
27	72.4	0.7	7446	8	ACA19708
28	72.4	0.7	7446	8	ACA19708
29	72.4	0.7	7446	10	ACF05848
30	68.8	0.7	7446	13	ADU74046
31	68.8	0.7	1998	3	AA170212
32	68.2	0.7	4911	8	ABT15076
33	68.2	0.7	4997	8	ACA47636
34	67.6	0.7	4590	1	AA160472
35	67.6	0.7	4590	1	AA160472
36	66.2	0.7	15507	8	AA153680
37	65.8	0.7	1000	12	ADQ62833
38	65.8	0.7	1000	12	ADQ62832
39	65.4	0.7	5940	3	AA170105
40	65.2	0.7	2289	14	AD272266
41	64.4	0.7	1712	6	ABK50884
42	63.4	0.7	612	14	ACL64894
43	63.2	0.6	2407	8	ADA89805
44	63.2	0.6	3095	2	AAQ03875
45	63.2	0.6	6644	2	AA133181

## ALIGNMENTS

RESULT 1	AA167640	standard; DNA; 9726 BP.
ID	AA167640	standard; DNA; 9726 BP.
XX	AA167640;	
AC	AA167640;	
XX	AA167640;	
DT	27-FEB-2002	(first entry)
XX	27-FEB-2002	(first entry)
DE	F. necrophorum leukotoxin protein encoding DNA.	
XX	F. necrophorum leukotoxin protein encoding DNA.	
KW	Leukotoxin; infection; immunisation; liver abscess; foot rot; bacteriophage; vaccine; ds.	
XX	Leukotoxin; infection; immunisation; liver abscess; foot rot; bacteriophage; vaccine; ds.	
OS	Fusobacterium necrophorum.	
XX	Fusobacterium necrophorum.	
FH	Key	Location/Qualifiers
FT	CDS	1..9726
FT		/tag= a
FT		/product= "Leukotoxin"
XX	WO200180886-A2.	
PN	WO200180886-A2.	
XX	WO200180886-A2.	
PD	01-NOV-2001.	
XX	01-NOV-2001.	
XX	25-APR-2001; 2001WO-US013240.	
PF	25-APR-2001; 2001WO-US013240.	
XX	25-APR-2001; 2001WO-US013240.	
PR	25-APR-2001; 2001WO-US013240.	
XX	25-APR-2001; 2001WO-US013240.	
PR	24-APR-2001; 2001US-00841786.	
XX	24-APR-2001; 2001US-00841786.	
PA	(UNITV ) UNIV KANSAS STATE RES FOUND.	
XX	(UNITV ) UNIV KANSAS STATE RES FOUND.	
XX	Nagaraja TG, Stewart GC, Narayanan SK, Chengappa MM;	
PI	Nagaraja TG, Stewart GC, Narayanan SK, Chengappa MM;	
DR	WPI; 2002-049245/06.	
XX	WPI; 2002-049245/06.	
XX	P-FSDB; AAG66005.	
XX	P-FSDB; AAG66005.	
PT	Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an	
XX	Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an	
PT	animal against an infection e.g. foot rot, or liver abscesses caused by	
XX	animal against an infection e.g. foot rot, or liver abscesses caused by	
XX	the bacterium.	
XX	the bacterium.	
PS	Claim 10; Page 90-95; 108bp; English.	
XX	Claim 10; Page 90-95; 108bp; English.	

13

CC The invention provides an isolated *Fusobacterium necrophorum* leukotoxin  
CC polypeptide and its truncated versions. The leukotoxin protein can be  
CC expressed by standard recombinant methodology. The leukotoxin gene is  
CC useful for preparing a vaccine which confers effective immunity against  
CC infection caused by *F. necrophorum*. The vaccine is useful for immunising  
CC an animal against liver abscesses caused by *F. necrophorum* and for  
CC preventing foot rot caused by *F. necrophorum* infection. The present  
CC sequence represents a DNA encoding a *F. necrophorum* full-length  
CC leukotoxin polypeptide  
XX

Sequence 9726 BP; 3678 A; 1304 C; 2238 G; 2506 T; 0 U; 0 Other;

Query Match 100.0%; Score 9726; DB 6; Length 9726;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGACGGGATCAAAATTAACGTTCCAGAGACAGAGAGAGATATCAGATTCTAAAAA 60
DB 1 ATGACGGGATCAAAATTAACGTTCCAGAGACAGAGAGAGATATCAGATTCTAAAAA 60
QY 61 GTTTTAATGATTTGGGATTTGATTTACACTATGACGGGTAAATGATCAATC 120
DB 61 GTTTTAATGATTTGGGATTTGATTTACACTATGACGGGTAAATGATCAATC 120
QY 121 ACCGGGACTGAGAAATTTGGAACAAAATAGAAAAAAGATATGTTATGACATTA 180
DB 121 ACCGGGACTGAGAAATTTGGAACAAAATAGAAAAAAGATATGTTATGACATTA 180
QY 181 ACCGGAATGAGAAATTTGGAACAAAATAGAAAAAAGATATGTTATGACATTA 180
DB 181 ACCGGAATGAGAAATTTGGAACAAAATAGAAAAAAGATATGTTATGACATTA 180
QY 181 ACAAACAAGATTCAGAGGAGAACGCTTTTACAGTTTAAATGATTTGCTTTAA 240
DB 181 ACAAACAAGATTCAGAGGAGAACGCTTTTACAGTTTAAATGATTTGCTTTAA 240
QY 241 AATAATATGCAATCTATATTTTGGGAAAAAGATAGACGGGGTAAATATCTTTT 300
DB 241 AATAATATGCAATCTATATTTTGGGAAAAAGATAGACGGGGTAAATATCTTTT 300
QY 301 AACTTGTCAATGAGAAAAATGAGTAGATGGATTTATCAACGGAATTCAGAAAA 360
DB 301 AACTTGTCAATGAGAAAAATGAGTAGATGGATTTATCAACGGAATTCAGAAAA 360
QY 361 ATTGAGAGAAATTTATTTCTTAAGCTCGAAAGGATGCGTAGAAAAATGAGTT 420
DB 361 ATTGAGAGAAATTTATTTCTTAAGCTCGAAAGGATGCGTAGAAAAATGAGTT 420
QY 421 ATCAATGCTGTTCTTTTCAATCTATTTTCCAAAAACAAGATTTTAAAGAGCT 480
DB 421 ATCAATGCTGTTCTTTTCAATCTATTTTCCAAAAACAAGATTTTAAAGAGCT 480
QY 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGAAATCATTCAGTAGATGAAAA 540
DB 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGAAATCATTCAGTAGATGAAAA 540
QY 541 ATTCAATGAAATCCGAATGGAAGCATTCAGTAGAAGAAAAATCAATGCTTT 600
DB 541 ATTCAATGAAATCCGAATGGAAGCATTCAGTAGAAGAAAAATCAATGCTTT 600
QY 601 ATCCGTTTATATGCGCGGATATTTAGATTGAAGATCACTGAATCAAGAG 660
DB 601 ATCCGTTTATATGCGCGGATATTTAGATTGAAGATCACTGAATCAAGAG 660
QY 661 ACAGATTTTAAAAATTTAGTCAATATTAAGTGAATTAATTTCTGCTGAC 720
DB 661 ACAGATTTTAAAAATTTAGTCAATATTAAGTGAATTAATTTCTGCTGAC 720
QY 721 TTTAAAGCTACCAAGCAAAATCTGAGATATTTATTTCTTCACTCAATAT 780
DB 721 TTTAAAGCTACCAAGCAAAATCTGAGATATTTATTTCTTCACTCAATAT 780
QY 781 CAAAAGCTATGGAAGAAAAATCACTGTTGAAAGAGATGAAAGAAATGTA 840
DB 781 CAAAAGCTATGGAAGAAAAATCACTGTTGAAAGAGATGAAAGAAATGTA 840
QY 841 AATACCAAGCAAAATATGAAATCTGATGCTGTTGAAAGAGATGAAATTA 900
```

```
DB 841 AATACCAAGCAAAATATGAAATCTGATCTGTAATGGAAGCAGATGAAAAATTA 900
QY 901 AGTGGAAAGCTACCAATGAGAGATTTTAAAGAAAGAGGAAAAAGAACTTA 960
DB 901 AGTGGAAAGCTACCAATGAGAGATTTTAAAGAAAGAGGAAAAAGAACTTA 960
QY 961 ACTCCTTAATGTTATCAGATGAGAGCTCCGTAAGATTAATTAAGAAAGCTTA 1020
DB 961 ACTCCTTAATGTTATCAGATGAGAGCTCCGTAAGATTAATTAAGAAAGCTTA 1020
QY 1021 GGAAGAAATGTTGACATTTACAGCTGAGCAAGAAATTTCTATGATGCAATTA 1080
DB 1021 GGAAGAAATGTTGACATTTACAGCTGAGCAAGAAATTTCTATGATGCAATTA 1080
QY 1081 AAGCTTGCAAAAGCACTCTTTTACGTTTTCAGGTTCTATTTCTCTATCAATTA 1140
DB 1081 AAGCTTGCAAAAGCACTCTTTTACGTTTTCAGGTTCTATTTCTCTATCAATTA 1140
QY 1141 GGAATTTTAAAGTTTATGACAAAGTAAAGTCAAGTGTGTTAATGAAAGATGCC 1200
DB 1141 GGAATTTTAAAGTTTATGACAAAGTAAAGTCAAGTGTGTTAATGAAAGATGCC 1200
QY 1201 GAAGCAAGAAAGAAAGCAAAATATTTCAATCTTACAGTGAATGAGCACTAG 1260
DB 1201 GAAGCAAGAAAGAAAGCAAAATATTTCAATCTTACAGTGAATGAGCACTAG 1260
QY 1261 GCACTACTTCTCCATTTAAATTTCAATTTATTTTGGAAAGCCAAATGAA 1320
DB 1261 GCACTACTTCTCCATTTAAATTTCAATTTATTTTGGAAAGCCAAATGAA 1320
QY 1321 CTCACTATCGAGCGGATATATTTCTGCAAAAGTAATTCGAATGTAATTA 1380
DB 1321 CTCACTATCGAGCGGATATATTTCTGCAAAAGTAATTCGAATGTAATTA 1380
QY 1381 GAAGTAAATCGAAGGAAAGCAGATATTTCAATCAATCTGAAATATCTAAT 1440
DB 1381 GAAGTAAATCGAAGGAAAGCAGATATTTCAATCAATCTGAAATATCTAAT 1440
QY 1441 TCTGTTTCTGTTGGAAGCATGAGATTTCAATTAAGTCTTTCAATTTG 1500
DB 1441 TCTGTTTCTGTTGGAAGCATGAGATTTCAATTAAGTCTTTCAATTTG 1500
QY 1501 GAAGGAAATTAATCTTCCGTCAAGATGCTAAAGAGCAAAAGTAAATCA 1560
DB 1501 GAAGGAAATTAATCTTCCGTCAAGATGCTAAAGAGCAAAAGTAAATCA 1560
QY 1561 GATGATGTAATGAGAAAGTGAAGCGATTAATTCATTCAGCTGCTGTAAG 1620
DB 1561 GATGATGTAATGAGAAAGTGAAGCGATTAATTCATTCAGCTGCTGTAAG 1620
QY 1621 TTTGGGGAATAGTGGTATGAGGCTTGTGCTGCAATATTTCTAATTAAT 1680
DB 1621 TTTGGGGAATAGTGGTATGAGGCTTGTGCTGCAATATTTCTAATTAAT 1680
QY 1681 CGTATAGATGTAATGATATCTACATGCAAGCAAGAGCACTAAATGTAAG 1740
DB 1681 CGTATAGATGTAATGATATCTACATGCAAGCAAGAGCACTAAATGTAAG 1740
QY 1741 ATTAATAAAATAGTGTCTGCAAAACAGATCTGATTTGGAACTTCCAA 1800
DB 1741 ATTAATAAAATAGTGTCTGCAAAACAGATCTGATTTGGAACTTCCAA 1800
QY 1801 GATCAGCTTTATGAAATCAGGTCATCTAAATCAATTTTAAATGCAATTA 1860
DB 1801 GATCAGCTTTATGAAATCAGGTCATCTAAATCAATTTTAAATGCAATTA 1860
QY 1861 GAGAGAGCAGTGTCAATGAGAAATTAAGAAATTAAGCACTTAATTAAG 1920
DB 1861 GAGAGAGCAGTGTCAATGAGAAATTAAGAAATTAAGCACTTAATTAAG 1920
QY 1921 GTGCTGCAACCATAGCAAAATCATTAATCTGCTTGTGCAATGAGAGAGT 1980
```

Db 1921 GTGCTGCACCATGACAAATCATATATCTCTGCTGCGCAATAGAGAGTGA 1980  
Qy 1981 AGACTTTCTTCAGAGTGGAAAGGAGTATATGTAAGGCAATTAAATGAAGCTCAAAATCTT 2040  
Db 1981 AGACTTTCTTCAGAGTGGAAAGGAGTATATGTAAGGCAATTAAATGAAGCTCAAAATCTT 2040  
Qy 2041 CGAGGACCTACGTCAGTGAAGTGGCTGTACGAAAGAAAGAAAGAAAGAAAGAAAGCTTAT 2100  
Db 2041 CGAGGACCTACGTCAGTGAAGTGGCTGTACGAAAGAAAGAAAGAAAGAAAGAAAGCTTAT 2100  
Qy 2101 GGAATATGACGAGTTTTTATATGAACTATATATATATATATATATATATATATATATAT 2160  
Db 2101 GGAATATGACGAGTTTTTATATGAACTATATATATATATATATATATATATATATATAT 2160  
Qy 2161 CATGCTGAATGGTATCGGAAGGAAATTTGATATCAACAGTGAATATATATATATATATAT 2220  
Db 2161 CATGCTGAATGGTATCGGAAGGAAATTTGATATCAACAGTGAATATATATATATATATAT 2220  
Qy 2221 AAAATTCCTTCAAAAATGGCAAAAGTCTGTATATGATTAATTAAGAGAGCT 2280  
Db 2221 AAAATTCCTTCAAAAATGGCAAAAGTCTGTATATGATTAATTAAGAGAGCT 2280  
Qy 2281 TTTGGAAAAAGAAACGAAATCTCAGATATGATCCGAAAGATATGAAATCTATGAAAA 2340  
Db 2281 TTTGGAAAAAGAAACGAAATCTCAGATATGATCCGAAAGATATGAAATCTATGAAAA 2340  
Qy 2341 TTTATGAATGCAATTTTCAGAAAAATTTGATGAAAAACCGAGCTTTTACTAATGCTGA 2400  
Db 2341 TTTATGAATGCAATTTTCAGAAAAATTTGATGAAAAACCGAGCTTTTACTAATGCTGA 2400  
Qy 2401 AGAATGACAAATTAATCTCCGATGGAATTTCAAAAACGGAACGCTATATGAATTTGA 2460  
Db 2401 AGAATGACAAATTAATCTCCGATGGAATTTCAAAAACGGAACGCTATATGAATTTGA 2460  
Qy 2461 AACTATGTTCAAGGAGAAATGAAAAATTAGAGAAAAATTAACCGAAAGATTTAAAGCT 2520  
Db 2461 AACTATGTTCAAGGAGAAATGAAAAATTAGAGAAAAATTAACCGAAAGATTTAAAGCT 2520  
Qy 2521 TTTTCAGAGAGATGAGTGAATTAAGAACTTTGAATTTTACAGAGTGAAGAAAT 2580  
Db 2521 TTTTCAGAGAGATGAGTGAATTAAGAACTTTGAATTTTACAGAGTGAAGAAAT 2580  
Qy 2581 TATGCAAAATTTTCACTTTTACCTTCCGAGCTTAATGGAAGAAAGATGTTCTTCT 2640  
Db 2581 TATGCAAAATTTTCACTTTTACCTTCCGAGCTTAATGGAAGAAAGATGTTCTTCT 2640  
Qy 2641 GTGGGAGAGCTGTTTGTGGGTAGAACAGAAATTTATGCAAGGTATCCGTTGAAAA 2700  
Db 2641 GTGGGAGAGCTGTTTGTGGGTAGAACAGAAATTTATGCAAGGTATCCGTTGAAAA 2700  
Qy 2701 GGAGCTAAACTGTGCAAAAAAAGATTTAAATATTAAGACTATCAATTAAGCAGAAACA 2760  
Db 2701 GGAGCTAAACTGTGCAAAAAAAGATTTAAATATTAAGACTATCAATTAAGCAGAAACA 2760  
Qy 2761 GTGAATTTAGTTGAAAAATTTGACTTGGAAGAGCATCCGAAAGTGCAGTCCGA 2820  
Db 2761 GTGAATTTAGTTGAAAAATTTGACTTGGAAGAGCATCCGAAAGTGCAGTCCGA 2820  
Qy 2821 GGAAGATTAAATGTTCAAAAGATGAAAAATTCAGCTATGCTGAAGCTTAAAGAAAGCT 2880  
Db 2821 GGAAGATTAAATGTTCAAAAGATGAAAAATTCAGCTATGCTGAAGCTTAAAGAAAGCT 2880  
Qy 2881 GAATTAATCAGAGAAAAATTAATGAGATGCAATTAAGCAGACTTTTCAATGAGCGGA 2940  
Db 2881 GAATTAATCAGAGAAAAATTAATGAGATGCAATTAAGCAGACTTTTCAATGAGCGGA 2940  
Qy 2941 TCTTTTAAATGCTGCTCAGGTGGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGGT 3000  
Db 2941 TCTTTTAAATGCTGCTCAGGTGGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGGT 3000  
Qy 3001 ATCAGTAAGGCAAGGTTTCCATGATGACGAAGCATATTTGAAAGCTAATTAATAAATTT 3060  
Db 3001 ATCAGTAAGGCAAGGTTTCCATGATGACGAAGCATATTTGAAAGCTAATTAATAAATTT 3060

Qy 3061 GCTTTAAACAGTAAGAAATGATTAATCTTGTGTAATGCAATGCTCCGGTTACGCGGAATCCGA 3120  
Db 3061 GCTTTAAACAGTAAGAAATGATTAATCTTGTGTAATGCAATGCTCCGGTTACGCGGAATCCGA 3120  
Qy 3121 ACGAAAAATGCGCGGCTGCGGGGTGCTGTTCCGTAATATGATATATTTCAACAAA 3180  
Db 3121 ACGAAAAATGCGCGGCTGCGGGGTGCTGTTCCGTAATATGATATATTTCAACAAA 3180  
Qy 3181 GCTTCCATTTGAAGATTAATGACGAAGAACAAATGTAATTAAGATTAAGATGATGA 3240  
Db 3181 GCTTCCATTTGAAGATTAATGACGAAGAACAAATGTAATTAAGATTAAGATGATGA 3240  
Qy 3241 GTAACAGTACCTGCGGAATCTTTAGAAATGATGCAAAAAACGACCGAACATCAACAGT 3300  
Db 3241 GTAACAGTACCTGCGGAATCTTTAGAAATGATGCAAAAAACGACCGAACATCAACAGT 3300  
Qy 3301 ATTTCTGTTGCGCGAGAAATTAATAGGTTGGAAGTAAACCGAGTGAAGAAAAACGAAA 3360  
Db 3301 ATTTCTGTTGCGCGAGAAATTAATAGGTTGGAAGTAAACCGAGTGAAGAAAAACGAAA 3360  
Qy 3361 TCAGAAAGAAACCCAGAGGATTTTGGCAAAATCCGAAACAAAGTGAATCTGTAAA 3420  
Db 3361 TCAGAAAGAAACCCAGAGGATTTTGGCAAAATCCGAAACAAAGTGAATCTGTAAA 3420  
Qy 3421 AATTAATTTAGGATGATGATGATTCATTAACAGAAAAATTAACAATTAATTAATTTCTGA 3480  
Db 3421 AATTAATTTAGGATGATGATGATTCATTAACAGAAAAATTAACAATTAATTAATTTCTGA 3480  
Qy 3481 GGAGTAAATAAAGCGCGGAATCTTCTCGAATGAAATGAAATGAAATGAAATGAAATGAA 3540  
Db 3481 GGAGTAAATAAAGCGCGGAATCTTCTCGAATGAAATGAAATGAAATGAAATGAAATGAA 3540  
Qy 3541 TCTTCAAGTTGGAGCTTTCTGGAAGTGTTCCTTCAATTAATTAATTAATTAATTAATTTCT 3600  
Db 3541 TCTTCAAGTTGGAGCTTTCTGGAAGTGTTCCTTCAATTAATTAATTAATTAATTAATTTCT 3600  
Qy 3601 GCTGTGATGATGAGTAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
Db 3601 GCTGTGATGATGAGTAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
Qy 3661 TCTTCTGATTTCACTTTTGTGAGACATGCGGCGGATCTGCTGCACTTCAATGATCAT 3720  
Db 3661 TCTTCTGATTTCACTTTTGTGAGACATGCGGCGGATCTGCTGCACTTCAATGATCAT 3720  
Qy 3721 ATTGAAGTGAATATGCAACATCACTGCTGTTTACTGGAAGCGGCTGCTGTAATAT 3780  
Db 3721 ATTGAAGTGAATATGCAACATCACTGCTGTTTACTGGAAGCGGCTGCTGTAATAT 3780  
Qy 3781 ATTCAAGTAAACAAGTCTTTGGTTAAAAATTAATTAATTAATTAATTAATTAATTAAT 3840  
Db 3781 ATTCAAGTAAACAAGTCTTTGGTTAAAAATTAATTAATTAATTAATTAATTAATTAAT 3840  
Qy 3841 AAAGTAAATGCTTGTGAGAGAACTCAAGTATGCAAGTATGCAATTAATTAATTAATTAAT 3900  
Db 3841 AAAGTAAATGCTTGTGAGAGAACTCAAGTATGCAAGTATGCAATTAATTAATTAATTAAT 3900  
Qy 3901 AAAGAAAGTGAAGCAAGGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960  
Db 3901 AAAGAAAGTGAAGCAAGGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960  
Qy 3961 GTGAACAAATGAAGTTTCTGCAAAATCAAGAAATTAATTAATTAATTAATTAATTAATTAAT 4020  
Db 3961 GTGAACAAATGAAGTTTCTGCAAAATCAAGAAATTAATTAATTAATTAATTAATTAATTAAT 4020  
Qy 4021 CAAAAAATGATGTTGATGCTCACTGCTTATCAAGCGGACACCAAGTGAACAGAGCTTTTA 4080  
Db 4021 CAAAAAATGATGTTGATGCTCACTGCTTATCAAGCGGACACCAAGTGAACAGAGCTTTTA 4080  
Qy 4081 AATTTAACAAGCTGGAAGATCAATGGAATCTGTAAGGCTTATCTGTGCAATTAATTAATTAAT 4140  
Db 4081 AATTTAACAAGCTGGAAGATCAATGGAATCTGTAAGGCTTATCTGTGCAATTAATTAATTAAT 4140

QY 4141 AACAAAGTAATGCTTCTATTAGTGTGGAGATATACTAACGTTAATGACGGAC 4200  
DB 4141 AACAAAGTAATGCTTCTATTAGTGTGGAGATATACTAACGTTAATGACGGAC 4200  
QY 4201 GCAAAAGCTCTTTAGCAACCACTCAAGTGACTGCTGCACTGACGACGAGGACAAATT 4260  
DB 4201 GCAAAAGCTCTTTAGCAACCACTCAAGTGACTGCTGCACTGACGAGGACAAATT 4260  
QY 4261 AGTTCTGAGCGGAGTTAGAAATTATCAAGGGGCTGTTCTGCAATTAAGTTGACAT 4320  
DB 4261 AGTTCTGAGCGGAGTTAGAAATTATCAAGGGGCTGTTCTGCAATTAAGTTGACAT 4320  
QY 4321 GACGTGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAATGAATCAATGTCATT 4380  
DB 4321 GACGTGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAATGAATCAATGTCATT 4380  
QY 4381 GCCAAAGTGTCAAAAGAAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGA 4440  
DB 4381 GCCAAAGTGTCAAAAGAAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGA 4440  
QY 4441 AAAAGTAAAAATATTTAGAAAGATGCTGATTAATAAGACTGGAATGTTATTAACG 4500  
DB 4441 AAAAGTAAAAATATTTAGAAAGATGCTGATTAATAAGACTGGAATGTTATTAACG 4500  
QY 4501 AAGGAACACTAGAAAAAGCAAAAGAAAAAGAGAGCGTCAATGTAAATGCTGCTTTA 4560  
DB 4501 AAGGAACACTAGAAAAAGCAAAAGAAAAAGAGAGCGTCAATGTAAATGCTGCTTTA 4560  
QY 4561 TCGGTGCTGGAACGGATTAATCCGCTGAGAGAGTACGTAATTCAGTCAATCTGTTAAA 4620  
DB 4561 TCGGTGCTGGAACGGATTAATCCGCTGAGAGAGTACGTAATTCAGTCAATCTGTTAAA 4620  
QY 4621 AATAATTTAAAGCAGAAATGAGTGAAGCAATAAGAAAGCGGAGAGATTAATTAAT 4680  
DB 4621 AATAATTTAAAGCAGAAATGAGTGAAGCAATAAGAAAGCGGAGAGATTAATTAAT 4680  
QY 4681 GCGAAACATGTAAATGTGAGAGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGA 4740  
DB 4681 GCGAAACATGTAAATGTGAGAGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGA 4740  
QY 4741 CTTCGTATCAGCAAAAGATGCTTTTCAAGAAATGGGATCTGAGCACTGGCAAGACTTAACA 4800  
DB 4741 CTTCGTATCAGCAAAAGATGCTTTTCAAGAAATGGGATCTGAGCACTGGCAAGACTTAACA 4800  
QY 4801 AATGACAGATTTGCAAAAGGTGATAAGAGAAATTTCTGCTGATTCCTTAATGTGAAC 4860  
DB 4801 AATGACAGATTTGCAAAAGGTGATAAGAGAAATTTCTGCTGATTCCTTAATGTGAAC 4860  
QY 4861 GCAAAATAATTCATTTCTGGGGTGAATGTTGCGGAAACCAATTCGCGGTTCTTTCTACG 4920  
DB 4861 GCAAAATAATTCATTTCTGGGGTGAATGTTGCGGAAACCAATTCGCGGTTCTTTCTACG 4920  
QY 4921 GCGGTGAGAGCTGCTTTTGCATTAATACTCTTCAATAAACCCTGCTTGATTAACA 4980  
DB 4921 GCGGTGAGAGCTGCTTTTGCATTAATACTCTTCAATAAACCCTGCTTGATTAACA 4980  
QY 4981 GAAAGCAAGGTAAATCCTTTAGTGAAGAAAGTAAAGCAAGTCAAGTCAAGCTTTGAAT 5040  
DB 4981 GAAAGCAAGGTAAATCCTTTAGTGAAGAAAGTAAAGCAAGTCAAGTCAAGCTTTGAAT 5040  
QY 5041 GATTCCTCATATTACAAACGTTTCTGCTGAGCGCTGCAAGTATTAACAGGCTGGAATC 5100  
DB 5041 GATTCCTCATATTACAAACGTTTCTGCTGAGCGCTGCAAGTATTAACAGGCTGGAATC 5100  
QY 5101 GAGAGAAATGTATCTGTCAATCTGTGTTCTGATGAACGAAAGCTTTAGTATTAATCT 5160  
DB 5101 GAGAGAAATGTATCTGTCAATCTGTGTTCTGATGAACGAAAGCTTTAGTATTAATCT 5160  
QY 5161 GAGTTGAAGAGGTAAAGTTCTTCAATGTAAGTGAAGAAAGTCAAAAGCAATTAATCA 5220  
DB 5161 GAGTTGAAGAGGTAAAGTTCTTCAATGTAAGTGAAGAAAGTCAAAAGCAATTAATCA 5220  
QY 5221 ATTGCCGAAATGCAAAATGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACA 5280

DB 5221 ATTGCCGAAATGCAAAATGAGAAAGCGGCTGAGTTGAGCAACAGTTGCTCATACA 5280  
QY 5281 AATATTGAAAAACAATCACTGTTATAGCTTTTGTAAAAAACAGTAAATTAACAAGCGCAAT 5340  
DB 5281 AATATTGAAAAACAATCACTGTTATAGCTTTTGTAAAAAACAGTAAATTAACAAGCGCAAT 5340  
QY 5341 GATCAAGATAGAAAAAATATCATATGTGACTCAAAAAGATTATATAGCAACCAATCACTTA 5400  
DB 5341 GATCAAGATAGAAAAAATATCATATGTGACTCAAAAAGATTATATAGCAACCAATCACTTA 5400  
QY 5401 GCAGTCGAGTTGAGAGAGCAAAAGAGCTCTGTGCAAGAGAGCTTCTGCAAGTACTAC 5460  
DB 5401 GCAGTCGAGTTGAGAGAGCAAAAGAGCTCTGTGCAAGAGAGCTTCTGCAAGTACTAC 5460  
QY 5461 TTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTAGAGGA 5520  
DB 5461 TTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTAGAGGA 5520  
QY 5521 GAAATTAATGGAATTAAGAAAGCAATGTTAATGTTCTAGAGTGAATTAACGAGTCA 5580  
DB 5521 GAAATTAATGGAATTAAGAAAGCAATGTTAATGTTCTAGAGTGAATTAACGAGTCA 5580  
QY 5581 GTGTCACAAAATGCGACACAGTCTTCCGAGCAAGTGAACAAGCTGCAAGTGAAGCTGGA 5640  
DB 5581 GTGTCACAAAATGCGACACAGTCTTCCGAGCAAGTGAACAAGCTGCAAGTGAAGCTGGA 5640  
QY 5641 GTACAGCTTAATTAATTAATTAACAACAATTCCTTCTGACATATTAATAAATAGTACTAAAT 5700  
DB 5641 GTACAGCTTAATTAATTAATTAACAACAATTCCTTCTGACATATTAATAAATAGTACTAAAT 5700  
QY 5701 GTACGAAATGCTTTGGTGTAAAAAGCAAAATCTGATTCATATTAATAAACAATTGAAATTGGA 5760  
DB 5701 GTACGAAATGCTTTGGTGTAAAAAGCAAAATCTGATTCATATTAATAAACAATTGAAATTGGA 5760  
QY 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGTGAACAGTTCTGTAGCAGTGAATTAAGATTGTA 5820  
DB 5761 GCTGAGTTGAGAGAGCTGAGAGAGCTGAGTGAACAGTTCTGTAGCAGTGAATTAAGATTGTA 5820  
QY 5821 AATAATACAGATACAGAAATTAATCACTGCAAAAATCACTGGAAGGAAATGTGGAAT 5880  
DB 5821 AATAATACAGATACAGAAATTAATCACTGCAAAAATCACTGGAAGGAAATGTGGAAT 5880  
QY 5881 ATTACAGAGTGTATGCGGTAAATGCTTAATATGAGAGAAACAGTGTCTGAGGTGCGCT 5940  
DB 5881 ATTACAGAGTGTATGCGGTAAATGCTTAATATGAGAGAAACAGTGTCTGAGGTGCGCT 5940  
QY 5941 GCAGCAATAGAGGCTCAACCAAGTGTGAATGAATTAACAGATCTACAAAAGCAATATGTA 6000  
DB 5941 GCAGCAATAGAGGCTCAACCAAGTGTGAATGAATTAACAGATCTACAAAAGCAATATGTA 6000  
QY 6001 AAAAGTTTACAGATGTTGCTTAAAGAGAAACAGATGTTATTAATCACTCAAGGCA 6060  
DB 6001 AAAAGTTTACAGATGTTGCTTAAAGAGAAACAGATGTTATTAATCACTCAAGGCA 6060  
QY 6061 GTAGTAAGAGGTAGATTAAGTATTCAAAATCTTAATTAATTAAGAGAGCTTAATCAAA 6120  
DB 6061 GTAGTAAGAGGTAGATTAAGTATTCAAAATCTTAATTAATTAAGAGAGCTTAATCAAA 6120  
QY 6121 AAAAGAAAAATAGTAATTAAGAAAAAGATTGTTACCAATAGTTCAGTCACTCATCTTTA 6180  
DB 6121 AAAAGAAAAATAGTAATTAAGAAAAAGATTGTTACCAATAGTTCAGTCACTCATCTTTA 6180  
QY 6181 AAATCTTTATTGGAACATGCCGCTGTTCAAGACAGCCGAGTGGCAAGAACTGTTAAT 6240  
DB 6181 AAATCTTTATTGGAACATGCCGCTGTTCAAGACAGCCGAGTGGCAAGAACTGTTAAT 6240  
QY 6241 ATCAACAAGGTTTATGAGAAACAGAGCTCTTGAGAAAAATTCATATTAATTAAGCAAAA 6300  
DB 6241 ATCAACAAGGTTTATGAGAAACAGAGCTCTTGAGAAAAATTCATATTAATTAAGCAAAA 6300  
QY 6301 CATTATTCGTAAATCAGAGATTAACAGAAATTCATCGAGTATAGTGTCTGTGGT 6360

Dh 6301 CATTATCTGTAAATATCAGAGATTACAGCAATTCATTCGAGTAGTAGTTCGTGTGT 6360  
Qy 6361 GTTGTGGAATATGAGTAGAGAGCTTCTGTGATACCAATATTTATTAAGAAATACC 6420  
Dh 6361 GTTGTGGAATATGAGTAGAGAGCTTCTGTGATACCAATATTTATTAAGAAATACC 6420  
Qy 6421 AAGACAAGAGTTGAAAAAATCTACATGCTGATGAAAGGTTTGGAGAAAGAGCTGAAAT 6480  
Dh 6421 AAGACAAGAGTTGAAAAAATCTACATGCTGATGAAAGGTTTGGAGAAAGAGCTGAAAT 6480  
Qy 6481 ACAGCAGATTCTAACCAAGAAATTTCTCTTTTGGAGTCGAGTCGACAGCCGAGGTA 6540  
Dh 6481 ACAGCAGATTCTAACCAAGAAATTTCTCTTTTGGAGTCGAGTCGACAGCCGAGGTA 6540  
Qy 6541 GGAGCCGAGTCGACAGAAACCGTTTCCGTAAATCAATTTGACAGAAAGACGGAAGTAGAT 6600  
Dh 6541 GGAGCCGAGTCGACAGAAACCGTTTCCGTAAATCAATTTGACAGAAAGACGGAAGTAGAT 6600  
Qy 6601 GTGGAAGAAAGCAAAATTTTGGTAAAAAAAGCTGAGATTACAGCAAAACGTTATAGTTCT 6660  
Dh 6601 GTGGAAGAAAGCAAAATTTTGGTAAAAAAAGCTGAGATTACAGCAAAACGTTATAGTTCT 6660  
Qy 6661 GTTGCAATTTGAAATGCGCAGTCGAGTCGCTGCAAAAAGAGCTGAAATTTGAGCAGCA 6720  
Dh 6661 GTTGCAATTTGAAATGCGCAGTCGAGTCGCTGCAAAAAGAGCTGAAATTTGAGCAGCA 6720  
Qy 6721 GTGCGAGTTACCAAAAGATGATCAAAACAGAGACGAGAGTAATAATTTCTAAATTTATG 6780  
Dh 6721 GTGCGAGTTACCAAAAGATGATCAAAACAGAGACGAGAGTAATAATTTCTAAATTTATG 6780  
Qy 6781 ACTCGAAACAATTTGATGATCTAATAGCAGAAATGAGATAAATTCAGTATCTGGAATCCGT 6840  
Dh 6781 ACTCGAAACAATTTGATGATCTAATAGCAGAAATGAGATAAATTCAGTATCTGGAATCCGT 6840  
Qy 6841 TCAGCCGAGTCGAGAAATTTCTGACGCGGAGTATCTGAGTGGTTTCTGCAATATAT 6900  
Dh 6841 TCAGCCGAGTCGAGAAATTTCTGACGCGGAGTATCTGAGTGGTTTCTGCAATATAT 6900  
Qy 6901 GCAAAATAGAGTGAAGACAGATATCGATCATAGTACTTTACACTTTCTGATGATGTAAT 6960  
Dh 6901 GCAAAATAGAGTGAAGACAGATATCGATCATAGTACTTTACACTTTCTGATGATGTAAT 6960  
Qy 6961 GTTAAAGCTCTTAATTAATTTTCGAAATTCCTTGAACGCGGTGAGAGACCCGAGCTCT 7020  
Dh 6961 GTTAAAGCTCTTAATTAATTTTCGAAATTCCTTGAACGCGGTGAGAGACCCGAGCTCT 7020  
Qy 7021 GCAGCAGTTACCGGAGTGGTTTCTGTTAACAATAATAGTTCTGTAGTACGTCGAGTT 7080  
Dh 7021 GCAGCAGTTACCGGAGTGGTTTCTGTTAACAATAATAGTTCTGTAGTACGTCGAGTT 7080  
Qy 7081 CACAATTAACCTGATTTGACTTCCGTACAGAGAAAGTAATGTAAACGCGCAAAAGAGAA 7140  
Dh 7081 CACAATTAACCTGATTTGACTTCCGTACAGAGAAAGTAATGTAAACGCGCAAAAGAGAA 7140  
Qy 7141 AAAAAATTTAAGCAACACAGCAAAATGACAGAAATGCGAGACACCAATCCGAGGCAT 7200  
Dh 7141 AAAAAATTTAAGCAACACAGCAAAATGACAGAAATGCGAGACACCAATCCGAGGCAT 7200  
Qy 7201 GTCTTGATTAATTAATTTTGAACAGCTGTAGAAAGATAGAAAAATTTCTGAAGAAAAAG 7260  
Dh 7201 GTCTTGATTAATTAATTTTGAACAGCTGTAGAAAGATAGAAAAATTTCTGAAGAAAAAG 7260  
Qy 7261 ACAGAGTTTAAAAAATTGAGACGAAGTTAACAAAGAACAGATAAAAAGTAAATGAT 7320  
Dh 7261 ACAGAGTTTAAAAAATTGAGACGAAGTTAACAAAGAACAGATAAAAAGTAAATGAT 7320  
Qy 7321 GCTACGAAAAAAATCTTCAATCAGCAGGATTTCTACAGAGATCTTCTGTAAGAGCG 7380  
Dh 7321 GCTACGAAAAAAATCTTCAATCAGCAGGATTTCTACAGAGATCTTCTGTAAGAGCG 7380  
Qy 7381 GATAGAGAGATCTACTCAGGAGAGAAATTAAGCAATTTGTAAGCTTCTGATATATTT 7440  
Dh 7381 GATAGAGAGATCTACTCAGGAGAGAAATTAAGCAATTTGTAAGCTTCTGATATATTT 7440

Qy 7441 GGAAGAAATGATGATATTTACACAGAGACAGAAATTAATCATCTTACTGTTGTTG 7500  
Dh 7441 GGAAGAAATGATGATATTTACACAGAGACAGAAATTAATCATCTTACTGTTGTTG 7500  
Qy 7501 GGAATCGCAGGTCCTTCCGATCAGGAAACAGTGGCAGTTACAAATATTTAAAGAAAT 7560  
Dh 7501 GGAATCGCAGGTCCTTCCGATCAGGAAACAGTGGCAGTTACAAATATTTAAAGAAAT 7560  
Qy 7561 TCCGAGTTACTGTTGAAAAATTTCTTTTGTGAAGAGCTGAAAAAGTAAATTTAGATCG 7620  
Dh 7561 TCCGAGTTACTGTTGAAAAATTTCTTTTGTGAAGAGCTGAAAAAGTAAATTTAGATCG 7620  
Qy 7621 GATATTAAGGAAATGTTGCTTTTAAACAGCATATCAAGGCTGTAGAGCATTTGGAATA 7680  
Dh 7621 GATATTAAGGAAATGTTGCTTTTAAACAGCATATCAAGGCTGTAGAGCATTTGGAATA 7680  
Qy 7681 GGAGCTGCTATGCAAAATTAATTTCTAATGGAAGATCAAAATATCAGTATTAATAATCT 7740  
Dh 7681 GGAGCTGCTATGCAAAATTAATTTCTAATGGAAGATCAAAATATCAGTATTAATAATCT 7740  
Qy 7741 AAGCTATTAAGAAAAATATTTGATGTTATTTGTAAGATTAATCCGAATTTGAGCCGAA 7800  
Dh 7741 AAGCTATTAAGAAAAATATTTGATGTTATTTGTAAGATTAATCCGAATTTGAGCCGAA 7800  
Qy 7801 GCAAAAGATTAACCGTGAAGCCGTAGCTGCGGAGCATTAATCTCAAAAGCAAAAGAT 7860  
Dh 7801 GCAAAAGATTAACCGTGAAGCCGTAGCTGCGGAGCATTAATCTCAAAAGCAAAAGAT 7860  
Qy 7861 GAATGAATTCAGAGTTGAAATTTGAGAAAGATATTTCAATGAAGAAATATAGATACT 7920  
Dh 7861 GAATGAATTCAGAGTTGAAATTTGAGAAAGATATTTCAATGAAGAAATATAGATACT 7920  
Qy 7921 AGCCCTTCTAAAGAAATTTGAGAAAGATCAATGTCAAAAGTGGAAAAAGAAACAGAGTC 7980  
Dh 7921 AGCCCTTCTAAAGAAATTTGAGAAAGATCAATGTCAAAAGTGGAAAAAGAAACAGAGTC 7980  
Qy 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGAGTAGCAAGGGGAGAAATTTATCCGAA 8040  
Dh 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGAGTAGCAAGGGGAGAAATTTATCCGAA 8040  
Qy 8041 GCAAAAGATGCGGAGAGCTCTTAATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100  
Dh 8041 GCAAAAGATGCGGAGAGCTCTTAATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100  
Qy 8101 CATGCAAGTAAATGTGAATATGGAACCAACATAAATGAAAGTAAACAGCAGTTCTAAA 8160  
Dh 8101 CATGCAAGTAAATGTGAATATGGAACCAACATAAATGAAAGTAAACAGCAGTTCTAAA 8160  
Qy 8161 GCAAGTAAACAGTTCTGTATTTGGAGAGAGTTGAGTCAACCAAGGCAAGACTCTGCTGCA 8220  
Dh 8161 GCAAGTAAACAGTTCTGTATTTGGAGAGAGTTGAGTCAACCAAGGCAAGACTCTGCTGCA 8220  
Qy 8221 GGTAAAACTATGATGTAAGTTGAGAGAAATTTGTTGAGAAACAATGATTTGATGCA 8280  
Dh 8221 GGTAAAACTATGATGTAAGTTGAGAGAAATTTGTTGAGAAACAATGATTTGATGCA 8280  
Qy 8281 ATTTCTAAAGTGAAGGTTTGGATGAGTAAAGTAACTGCTAAATCTTCTGTAGTATCA 8340  
Dh 8281 ATTTCTAAAGTGAAGGTTTGGATGAGTAAAGTAACTGCTAAATCTTCTGTAGTATCA 8340  
Qy 8341 GGAATGAGAGAGAAATTTCCCGGACAGAGAGTAACTTTCTACAGCAAAAGTAAATCT 8400  
Dh 8341 GGAATGAGAGAGAAATTTCCCGGACAGAGAGTAACTTTCTACAGCAAAAGTAAATCT 8400  
Qy 8401 GAATCCGTAGTTCGTTTACGAAAGCAAGTATTAAGAAATTAATGATTAACCAAAAAATAT 8460  
Dh 8401 GAATCCGTAGTTCGTTTACGAAAGCAAGTATTAAGAAATTAATGATTAACCAAAAAATAT 8460  
Qy 8461 ATTTGAGAGTCAATGCTCTTCTTTAAATGATACAAAGATGAGCGGAATATAGATCT 8520  
Dh 8461 ATTTGAGAGTCAATGCTCTTCTTTAAATGATACAAAGATGAGCGGAATATAGATCT 8520

```

Qy 8521 TTAGCGTAGCCGGTGTGCATGACACAGGAACAAACAGCATTTACAGATCAAAACAG 8580
Db 8521 TTAGCGTAGCCGGTGTGCATGACACAGGAACAAACAGCATTTACAGATCAAAACAG 8580
Qy 8581 TTTAATCTTACACCTGTAAATGAGAGAAAAGTATCTCAACTTCGTGCAAAAAGCTTTGGCT 8640
Db 8581 TTTAATCTTACACCTGTAAATGAGAGAAAAGTATCTCAACTTCGTGCAAAAAGCTTTGGCT 8640
Qy 8641 AAAAAATGAAAAATTTAGAAAATGTAAGAGAACTGAGAGAGCTTTAGTGGAGCGGAAACA 8700
Db 8641 AAAAAATGAAAAATTTAGAAAATGTAAGAGAACTGAGAGAGCTTTAGTGGAGCGGAAACA 8700
Qy 8701 GCAGCCGTTGAAAAATTAATACAAAGAGTACTACAGAGAGATTTGGTTCAGAAAATTTGGAA 8760
Db 8701 GCAGCCGTTGAAAAATTAATACAAAGAGTACTACAGAGAGATTTGGTTCAGAAAATTTGGAA 8760
Qy 8761 ATTGAGATTAATTTGAAAACGATTTGCAAGAGATTAATAGATTTGTAAGTCAACGAGAC 8820
Db 8761 ATTGAGATTAATTTGAAAACGATTTGCAAGAGATTAATAGATTTGTAAGTCAACGAGAC 8820
Qy 8821 GGAACCAAGAGAGTCTTGTGCGAAAAGATGGTATTTCTGTGAAAAATTAACAATTTCAAGG 8880
Db 8821 GGAACCAAGAGAGTCTTGTGCGAAAAGATGGTATTTCTGTGAAAAATTAACAATTTCAAGG 8880
Qy 8881 GAAACAAAATCATCCATTGAAGATTAAGCCAGAAATTTGGAACCGGAAGTGAATGTA 8940
Db 8881 GAAACAAAATCATCCATTGAAGATTAAGCCAGAAATTTGGAACCGGAAGTGAATGTA 8940
Qy 8941 GATGCTTTGAATGAACCTTGATGATGATCTACAGAGAAAAGTGGTGGCTATGTGGAATT 9000
Db 8941 GATGCTTTGAATGAACCTTGATGATGATCTACAGAGAAAAGTGGTGGCTATGTGGAATT 9000
Qy 9001 GGTATTTGAAAAATGTTGATGTAATATATGATTAAGAAAAATGTGAAGCCAAAATTCGGA 9060
Db 9001 GGTATTTGAAAAATGTTGATGTAATATATGATTAAGAAAAATGTGAAGCCAAAATTCGGA 9060
Qy 9061 AGACATGCTATTTGTGAAGATCTAGTGAAGAACAGAAATTCAGACATTTACAGAGCAAAA 9120
Db 9061 AGACATGCTATTTGTGAAGATCTAGTGAAGAACAGAAATTCAGACATTTACAGAGCAAAA 9120
Qy 9121 GTAAATATTTCTTGGAAAAAGAGACGCTGCACTGCACTGCATATTCGATGATACACTT 9180
Db 9121 GTAAATATTTCTTGGAAAAAGAGACGCTGCACTGCACTGCATATTCGATGATACACTT 9180
Qy 9181 TCCAAATGAGATGATTAATAAAAAATTTGGCAAGAGATGCACTTCTCAATTAATAACC 9240
Db 9181 TCCAAATGAGATGATTAATAAAAAATTTGGCAAGAGATGCACTTCTCAATTAATAACC 9240
Qy 9241 AAAAAATCAAAAAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9300
Db 9241 AAAAAATCAAAAAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9300
Qy 9301 GGGGTGCTGAAGCAAGAGCTGCAAGAGCCAAAGGAGCAAGTTAGTGAAGATCAAAATA 9360
Db 9301 GGGGTGCTGAAGCAAGAGCTGCAAGAGCCAAAGGAGCAAGTTAGTGAAGATCAAAATA 9360
Qy 9361 AATAGAACTAATTAATGATGATTTAGCAGAGAAAAATTTAAAAACAGAGAAACATCAATGTA 9420
Db 9361 AATAGAACTAATTAATGATGATTTAGCAGAGAAAAATTTAAAAACAGAGAAACATCAATGTA 9420
Qy 9421 TATGCCGATATGATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480
Db 9421 TATGCCGATATGATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480
Qy 9481 GCCAAAATCATGCTGCAAGCTGCTTGGCACTGCACTATTTGAAAAAAAATGAAGTAAAA 9540
Db 9481 GCCAAAATCATGCTGCAAGCTGCTTGGCACTGCACTATTTGAAAAAAAATGAAGTAAAA 9540
Qy 9541 TTTAATTAATGAGATCCGAGAAATTTAAAAATATCTGGCAAGATTTGGAAAGGGAGCTAAT 9600
Db 9541 TTTAATTAATGAGATCCGAGAAATTTAAAAATATCTGGCAAGATTTGGAAAGGGAGCTAAT 9600
Qy 9601 AAAAAACGTGCGTAGATCTTAATCAGGTAGACTGATTAAGGATTAATATATCATGCGAT 9660

```

```

Db 9601 AAAAAACGTGCGTAGATCTTAATCAGGTAGACTGATTAAGGATTAATATATCATGCGAT 9660
Qy 9661 TCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGGAAA 9720
Db 9661 TCTTCTGAAAAAGCATACAAAAATTTGACATATCAATCAAAAGAGAGAAAAAGGGAAA 9720
Qy 9721 AAATGA 9726
Db 9721 AAATGA 9726

RESULT 2
AA167647
ID AA167647 standard; DNA; 11130 BP.
XX
AC AA167647;
XX
DT 27-FEB-2002 (first entry)
XX
DE F. necrophorum leukotoxin gene sequence.
XX
KW Leukotoxin; infection; immunisation; liver abscess; foot rot;
KW bactericide; vaccine; ds.
XX
OS Fusobacterium necrophorum.
XX
PN WO200180886-A2.
XX
PD 01-NOV-2001.
XX
PF 25-APR-2001; 2001WO-US013240.
XX
PR 25-APR-2000; 2000US-00558257.
XX
PR 24-APR-2001; 2001US-00841786.
XX
PA (UNIV ) UNIV KANSAS STATE RES FOUND.
XX
PI Nagareja TG, Stewart GC, Narayanan SK, Chengappa MM;
XX
WP1: 2002-049245/06.
XX
DR
XX
PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an
PT animal against an infection e.g. foot rot, or liver abscesses caused by
PT the bacterium.
XX
PS Disclosure; Page 102-107; 108pp; English.
XX
CC The invention provides an isolated Fusobacterium necrophorum leukotoxin
CC polypeptide and its truncated versions. The leukotoxin protein can be
CC expressed by standard recombinant methodology. The leukotoxin gene is
CC useful for preparing a vaccine which confers effective immunity against
CC infection caused by F. necrophorum. The vaccine is useful for immunising
CC an animal against liver abscesses caused by F. necrophorum and for
CC preventing foot rot caused by F. necrophorum infection. The present
CC sequence represents the F. necrophorum full-length leukotoxin gene
CC sequence
XX
SQ Sequence 11130 BP; 4205 A; 1472 C; 2511 G; 2942 T; 0 U; 0 Other;
XX

Query Match 99.9%; Score 9718; DB 6; Length 11130;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 9721; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ATGAGCGGCATCAAAAATTAACGTTCAAGAGCAAGAGAAAGAGATATCAATTTCAAAAA 60
Db 1034 ATGAGCGGCATCAAAAATTAACGTTCAAGAGCAAGAGAAAGAGATATCAATTTCAAAAA 1093
Qy 61 GTTTTAATGATTTTGGATTTGTTGATTAACATGACTGACCGTGAAGGCTTAATGCAATC 120
Db 1094 GTTTTAATGATTTTGGATTTGTTGATTAACATGACTGACCGTGAAGGCTTAATGCAATC 1153
Qy 121 ACCGCACTGAGAAATTTTGAACAAAAATAGAAAAAAGATTAATGTTTATGACATTACT 180

```

```
1154 ACCGCACTGAGAAATTTTGGAAACAAATAATGAAAAAAGATTAAGTTTATGACATTAAT 1213
181 ACMAACAAATTCAGAGGGAGAACGCTTTTAAACGTTTAAATAGATTTGCTTTAAACAGAA 240
1214 ACMAACAAATTCAGAGGGAGAACGCTTTTAAACGTTTAAATAGATTTGCTTTAAACAGAA 1273
241 AATAATATGCAAAATCTATATTTTGGGAAAAAGATAGACGGGGTAAATATCTTTT 300
1274 AATAATATGCAAAATCTATATTTTGGGAAAAAGATAGACGGGGTAAATATCTTTT 1333
301 AACTTTGCAATGAGAAAAATGGAAGTAGATGGATTATCAAGGAAATTCAGAAAAATAAA 360
1334 AACTTTGCAATGAGAAAAATGGAAGTAGATGGATTATCAAGGAAATTCAGAAAAATAAA 1393
361 ATTGGAGGAAATTTATATTTCTTAAGCTCGAAAGGGATGCGATAGAAAAATGAGATT 420
1394 ATTGGAGGAAATTTATATTTCTTAAGCTCGAAAGGGATGCGATAGAAAAATGAGATT 1453
421 ATCAAATGCTGTTCTTTTCAATTCATATTTCCAAAACAAGATTTTAAAGAGCTTTG 480
1454 ATCAAATGCTGTTCTTTTCAATTCATATTTCCAAAACAAGATTTTAAAGAGCTTTG 1513
481 GAAAGAACCAACATGCTAAAGTTTAAATGGAATGATCCAGTAGATGAAAAAGTAAAA 540
1514 GAAAGAACCAACATGCTAAAGTTTAAATGGAATGATCCAGTAGATGAAAAAGTAAAA 1573
541 ATTCATTTGAATCCGAATGGAAGCATTAACGTAGAGAAAAATCAATCTGTTGAAGGC 600
1574 ATTCATTTGAATCCGAATGGAAGCATTAACGTAGAGAAAAATCAATCTGTTGAAGGC 1633
601 ATCGGTTATATGCGCGGATATTTAGATTGAAGATATGCAATCTAAAGACAGAAAT 660
1634 ATCGGTTATATGCGCGGATATTTAGATTGAAGATATGCAATCTAAAGACAGAAAT 1693
661 ACAGATTTTAAAAATTTAGTCAATATTTAGATCGAATAAATCTGCTGACCGGAGAT 720
1694 ACAGATTTTAAAAATTTAGTCAATATTTAGATCGAATAAATCTGCTGACCGGAGAT 1753
721 TTAAAAAGCTACCAAGACAAAAATCTGAGATATTAATCTTTCAGCTCACAATAGTTCCT 780
1754 TTAAAAAGCTACCAAGACAAAAATCTGAGATATTAATCTTTCAGCTCACAATAGTTCCT 1813
781 CAAAAAGCTATGGAAGAAAAATTTCACTGTTGAAAGAAATAGAGATATGTTAAAAAG 840
1814 CAAAAAGCTATGGAAGAAAAATTTCACTGTTGAAAGAAATAGAGATATGTTAAAAAG 1873
841 AATACCAAGCAAAATTTGAATCTGATGCTGTATTTGGAAGCAGATGAAAAATTAATAAT 900
1874 AATACCAAGCAAAATTTGAATCTGATGCTGTATTTGGAAGCAGATGAAAAATTAATAAT 1933
901 AGTGCMAAGCTACAATGSGAGATTTATTAAGAAAAAGGGAAAAAGAACTTAATAC 960
1934 AGTGCMAAGCTACAATGSGAGATTTATTAAGAAAAAGGGAAAAAGAACTTAATAC 1993
961 ACTCTTTAAGTTTATCAGATGTGAAAGCTTCGTAAGAGTAATAAGAAAAAGTCAAT 1020
1994 ACTCTTTAAGTTTATCAGATGTGAAAGCTTCGTAAGAGTAATAAGAAAAAGTCAAT 2053
1021 GGAAGAAATTTGACATTAACGCTGAAGCAAAAGAAATTTCTATGATGAACTTTAGTTACT 1080
2054 GGAAGAAATTTGACATTAACGCTGAAGCAAAAGAAATTTCTATGATGAACTTTAGTTACT 2113
1081 AAGCTTGCAAGCACTCTTTAGCTTTGTAACAGTTCTATTTCTCTATCAATTTAAAT 1140
2114 AAGCTTGCAAGCACTCTTTAGCTTTGTAACAGTTCTATTTCTCTATCAATTTAAAT 2173
1141 GGAATTTTGAAGTTTATGACAAAGTACAGTGTCTTATTTGAAAAAGATGCCAAAGTC 1200
2174 GGAATTTTGAAGTTTATGACAAAGTACAGTGTCTTATTTGAAAAAGATGCCAAAGTC 2233
1201 GAAGCAACAGAGAAAGGCAAAATATTCATTTTACGAGTAGAAGATCAATGCGA 1260
2234 GAAGCAACAGAGAAAGGCAAAATATTCATTTTACGAGTAGAAGATCAATGCGA 2293
1261 GCACTACTTCTCCCTTAAAAATTTTCAATTTATTTGGAGAAAGCCAAATGGAACCTT 1320
2294 GCACTACTTCTCCCTTAAAAATTTTCAATTTATTTGGAGAAAGCCAAATGGAACCTT 2353
1321 CTCAGTATCGAGCGGATATATTTTCTGCAAAAAGTAATTCCAATGTAATGTAAGGA 1380
2354 CTCAGTATCGAGCGGATATATTTTCTGCAAAAAGTAATTCCAATGTAATGTAAGGA 2413
1381 GAAGTAAAAATCGAAGGAAAGACAGATTAATCTTCAAAAATCTGAAAATCTATGATGCT 1440
2414 GAAGTAAAAATCGAAGGAAAGACAGATTAATCTTCAAAAATCTGAAAATCTATGATGCT 2473
1441 TCTGTTTCTGTTGGAACGATGAGATTCGAATAAGTAGCTTCTGATATTTGGAGC 1500
2474 TCTGTTTCTGTTGGAACGATGAGATTCGAATAAGTAGCTTCTGATATTTGGAGC 2533
1501 GAAAGGAAAAATTAATCTTCCGTCAGAAATGCTTAAAGGACAAAGTAAATCAGAAACG 1560
2534 GAAAGGAAAAATTAATCTTCCGTCAGAAATGCTTAAAGGACAAAGTAAATCAGAAACG 2593
1561 GATGATGTAATGTAAGAAAGTGAAGCATTAATTCATTCGAGCTGCTGTAAGAGTGA 1620
2594 GATGATGTAATGTAAGAAAGTGAAGCATTAATTCATTCGAGCTGCTGTAAGAGTGA 2653
1621 TTGGGGATATGTTGTAATGGGGTTGGTGCAAATATTTCTTAACTATTAATGCTTCCTC 1680
2654 TTGGGGATATGTTGTAATGGGGTTGGTGCAAATATTTCTTAACTATTAATGCTTCCTC 2713
1681 CGTATGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1740
2714 CGTATGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 2773
1741 ATTAATTAATAATATGTTCTGCAAAACAGATCTGATTTGGAACTTCCAAGTTTATGAAT 1800
2774 ATTAATTAATAATATGTTCTGCAAAACAGATCTGATTTGGAACTTCCAAGTTTATGAAT 2833
1801 GATCAGTTTATGAAATCAGGTCATCTAAATCAATTTTGAAGCAATAAAACGCGGTTT 1860
2834 GATCAGTTTATGAAATCAGGTCATCTAAATCAATTTTGAAGCAATAAAACGCGGTTT 2893
1861 GGAAGGACAGTGTCAATGAGGAAATTAAGAAATTAAGCTAACGAACTTATTTAGTGTCT 1920
2894 GGAAGGACAGTGTCAATGAGGAAATTAAGAAATTAAGCTAACGAACTTATTTAGTGTCT 2953
1921 GTGTCTGCAACCATATGCAATATCAATAATTTCTGCTTGTGTGCAATAGAGAGATGGA 1980
2954 GTGTCTGCAACCATATGCAATATCAATAATTTCTGCTTGTGTGCAATAGAGAGATGGA 3013
1981 AGACTTTCTTCAAGAGTGAAGGAGTAATGTAAGGCAATTAATGAACTCAAAATCTT 2040
3014 AGACTTTCTTCAAGAGTGAAGGAGTAATGTAAGGCAATTAATGAACTCAAAATCTT 3073
2041 CGAGCGACATACGTAAGTGAAGTGTGCTGTGACGAAGGAAAGAAAAAGAACTTATTT 2100
3074 CGAGCGACATACGTAAGTGAAGTGTGCTGTGACGAAGGAAAGAAAAAGAACTTATTT 3133
2101 GGAATATGACAGAGTTTATTAAGAAATTAATAATGCTTCTGTGCAATTTGCGCAT 2160
3134 GGAATATGACAGAGTTTATTAAGAAATTAATAATGCTTCTGTGCAATTTGCGCAT 3193
2161 CATGCTGAATTTGCTATCGGAAGAAATTTGATATCAACGTAAGAAATTAATAATTTGAAT 2220
3194 CATGCTGAATTTGCTATCGGAAGAAATTTGATATCAACGTAAGAAATTAATAATTTGAAT 3253
2221 AAAAATCTCTCAAAAATTTGCAAAAGTGTGTTATGTAATTTAGAACTTTTAAAGAGCT 2280
3254 AAAAATCTCTCAAAAATTTGCAAAAGTGTGTTATGTAATTTAGAACTTTTAAAGAGCT 3313
2281 TTGGAAGAAAGAAACGAAATCTCCAGAAATATGATCCGAAGATATTTGAATCTATTTGAAA 2340
3314 TTGGAAGAAAGAAACGAAATCTCCAGAAATATGATCCGAAGATATTTGAATCTATTTGAAA 3373
```

QY 2341 TTATTGAATGCATTTTCAGAAAAATTGATGAAAAACCGAGCTTTTACTTAATGCTGAA 2400  
DB 3374 TTATTGAATGCATTTTCAGAAAAATTGATGAAAAACCGAGCTTTTACTTAATGCTGAA 3433  
QY 2401 AGAATGACAAATTATCTCCCGATGGAACCTTCAAAAAACGAACTGCTATAGAAATTGCA 2460  
DB 3434 AGAATGACAAATTATCTCCCGATGGAACCTTCAAAAAACGAACTGCTATAGAAATTGCA 3493  
QY 2461 AACTATGTTACAGGAGAAATGAAAAATTAGAGAAAAATTACCGAAAGATTTTAAGCT 2520  
DB 3494 AACTATGTTACAGGAGAAATGAAAAATTAGAGAAAAATTACCGAAAGATTTTAAGCT 3553  
QY 2521 TTTTCAGAAAGATTGATGAGTGAATTAAAGAACTTTGAATTTTACAGAGTAGAAAT 2580  
DB 3554 TTTTCAGAAAGATTGATGAGTGAATTAAAGAACTTTGAATTTTACAGAGTAGAAAT 3613  
QY 2581 TATGCAAAATTTTCAACATTTTACCTCTCCGAGCTAATGAGAAAGAGATTTCTTCT 2640  
DB 3614 TATGCAAAATTTTCAACATTTTACCTCTCCGAGCTAATGAGAAAGAGATTTCTTCT 3673  
QY 2641 GTGGAGAGAGCTGTTTCGTGGGTAGAACAGAGAAATTATAGCAAGGTATCCGTTGAAAA 2700  
DB 3674 GTGGAGAGAGCTGTTTCGTGGGTAGAACAGAGAAATTATAGCAAGGTATCCGTTGAAAA 3733  
QY 2701 GGAGCTAAACTTGCTGCAAAAAAAGATTTAAATATATAAAAGCTATCAATTAACAGAAAA 2760  
DB 3734 GGAGCTAAACTTGCTGCAAAAAAAGATTTAAATATATAAAAGCTATCAATTAACAGAAAA 3793  
QY 2761 GTGAATTTTATGTTGGAAATTTATGACCTTGGAGAACAGTACATCCGAAAGTGCATCGGA 2820  
DB 3794 GTGAATTTTATGTTGGAAATTTATGACCTTGGAGAACAGTACATCCGAAAGTGCATCGGA 3853  
QY 2821 GGAAGATTTAATGTTCAAAAGATCGAAAAATTCAGCTATGTAAGAGCTTAAGAAAAAGCT 2880  
DB 3854 GGAAGATTTAATGTTCAAAAGATCGAAAAATTCAGCTATGTAAGAGCTTAAGAAAAAGCT 3913  
QY 2881 GAATTTATCAGAGAAAAATTTATATGACATGATGAAACAGACTTTTTCATGTAGCGGA 2940  
DB 3914 GAATTTATCAGAGAAAAATTTATATGACATGATGAAACAGACTTTTTCATGTAGCGGA 3973  
QY 2941 TCTTTTAAATGCTGCTCAGGTGGGAATGCAATCAATGGAATGGAAAGTTATAGTGAAGT 3000  
DB 3974 TCTTTTAAATGCTGCTCAGGTGGGAATGCAATCAATGGAATGGAAAGTTATAGTGAAGT 4033  
QY 3001 ATCAGTAAGGCAAGAGTTCCATTGATGACGAAGCATTTTGAAGCTTAATAAAAAATT 3060  
DB 4034 ATCAGTAAGGCAAGAGTTCCATTGATGACGAAGCATTTTGAAGCTTAATAAAAAATT 4093  
QY 3061 GCTTTAAACAGTAAGATGATACCTTCTGTTTGAATGCTGCCGGTTCAAGCGGAAATCGGA 3120  
DB 4094 GCTTTAAACAGTAAGATGATACCTTCTGTTTGAATGCTGCCGGTTCAAGCGGAAATCGGA 4153  
QY 3121 ACGAAAAATGCGCGCGGTGCTGCTGCGGTGAATGATTAATGATTTTCAACAAA 3180  
DB 4154 ACGAAAAATGCGCGCGGTGCTGCTGCGGTGAATGATTAATGATTTTCAACAAA 4213  
QY 3181 GCTTCATTGAAGATTAATGACGAAGGACAAAGTAATATGATTAAGAAATTAAGATGAA 3240  
DB 4214 GCTTCATTGAAGATTAATGACGAAGGACAAAGTAATATGATTAAGAAATTAAGATGAA 4273  
QY 3241 GTAAAGTAATCTGCGGAATCTTTAGAGATGATGCAAAAAACGACCGAACATCAACAGT 3300  
DB 4274 GTAAAGTAATCTGCGGAATCTTTAGAGATGATGCAAAAAACGACCGAACATCAACAGT 4333  
QY 3301 ATTTCTGTTGCGGAGGAATTAATAAGTTGGAAGTAAACCGAGTAGAAAAACCGAAA 3360  
DB 4334 ATTTCTGTTGCGGAGGAATTAATAAGTTGGAAGTAAACCGAGTAGAAAAACCGAAA 4393  
QY 3361 TCAGAGAAAGAACAGAGAGATTTTGTGCAAAATTCGGAACAAAGTGCATCTGTAAAA 3420  
DB 4394 TCAGAGAAAGAACAGAGAGATTTTGTGCAAAATTCGGAACAAAGTGCATCTGTAAAA 4453

QY 3421 AATAAAATTCAGGATAGTATGATTCATTTAACAGAAAAATTACAAATTACATTTCTGAA 3480  
DB 4454 AATAAAATTCAGGATAGTATGATTCATTTAACAGAAAAATTACAAATTACATTTCTGAA 4513  
QY 3481 GGAATTAATAAAAGCGGGGAATCTTCCCTCGAAGCTTTCTCATCTCCGATTAAGACCG 3540  
DB 4514 GGAATTAATAAAAGCGGGGAATCTTCCCTCGAAGCTTTCTCATCTCCGATTAAGACCG 4573  
QY 3541 TCTTTCAGTTTGGAGCTTCTGGAAGTGTCTTCAATTAATTTAAAAAGAAACATCT 3600  
DB 4574 TCTTTCAGTTTGGAGCTTCTGGAAGTGTCTTCAATTAATTTAAAAAGAAACATCT 4633  
QY 3601 GCTGTCGTAGATGAGATTAAGATTAATTTGAAGGAGCAAAATAAAAAGTGAAGTGA 3660  
DB 4634 GCTGTCGTAGATGAGATTAAGATTAATTTGAAGGAGCAAAATAAAAAGTGAAGTGA 4693  
QY 3661 TCTTCTGATTTCTACTTTTGTGAGACATGGGGCGGATCTGCTGCACTTCAGTGAATCAT 3720  
DB 4694 TCTTCTGATTTCTACTTTTGTGAGACATGGGGCGGATCTGCTGCACTTCAGTGAATCAT 4753  
QY 3721 ATTGGAAGTGGAAATTAAGCAATCAGTGTGCTGTTAGCTGAGACGGCTGTATAATAT 3780  
DB 4754 ATTGGAAGTGGAAATTAAGCAATCAGTGTGCTGTTAGCTGAGACGGCTGTATAATAT 4813  
QY 3781 ATTCAAGTAAAAACAAGTCTTGTGTTAAAAATAGATATTTGAAATGCCAATTAATT 3840  
DB 4814 ATTCAAGTAAAAACAAGTCTTGTGTTAAAAATAGATATTTGAAATGCCAATTAATT 4873  
QY 3841 AAAGTAAATGCTTTGAGTGAAGAGAACTCAATTAAGAGAGAGCAAGTTTGGAGCAATT 3900  
DB 4874 AAAGTAAATGCTTTGAGTGAAGAGAACTCAATTAAGAGAGAGCAAGTTTGGAGCAATT 4933  
QY 3901 AAAGAAAGTGAAGCAAGAAAAAATGATATGTAATGGAACCTTCTCTATCACTTA 3960  
DB 4934 AAAGAAAGTGAAGCAAGAAAAAATGATATGTAATGGAACCTTCTCTATCACTTA 4993  
QY 3961 GTGAACAAATGAAGTTCTGCAAAATCAGAAAAATTAATCAGTAGCAGAGAAATCTGAAAGC 4020  
DB 4994 GTGAACAAATGAAGTTCTGCAAAATCAGAAAAATTAATCAGTAGCAGAGAAATCTGAAAGC 5053  
QY 4021 CAAAAAATGGAATGTTGATGTCACCTGCTTAATCAAGCGGACACCCAAAGTACAGAGCTTTA 4080  
DB 5054 CAAAAAATGGAATGTTGATGTCACCTGCTTAATCAAGCGGACACCCAAAGTACAGAGCTTTA 5113  
QY 4081 AATTTCAAGCTGGAAGATCAAAATGGAACCTGTAAGGGCTACTGTAAGCTTGTGCCAAATTA 4140  
DB 5114 AATTTCAAGCTGGAAGATCAAAATGGAACCTGTAAGGGCTACTGTAAGCTTGTGCCAAATTA 5173  
QY 4141 AACAACAAAATGATGCTTCTTATTAATGTTGGAGATATTAATCACTTAATTCGACCGAC 4200  
DB 5174 AACAACAAAATGATGCTTCTTATTAATGTTGGAGATATTAATCACTTAATTCGACCGAC 5233  
QY 4201 GCCAAAAGCTCTTTTATGCAACCATCTCAATGCTGTGCAAGTGAACGACGGAGGCAATTT 4260  
DB 5234 GCCAAAAGCTCTTTTATGCAACCATCTCAATGCTGTGCAAGTGAACGACGGAGGCAATTT 5293  
QY 4261 AGTTCTGAGCGGGATTAAGAAATTAATCAAGGGCTGTTTCTGTCAATTAAGATTGACAAAT 4320  
DB 5294 AGTTCTGAGCGGGATTAAGAAATTAATCAAGGGCTGTTTCTGTCAATTAAGATTGACAAAT 5353  
QY 4321 GACGTGAAGTGAAGTGAATGATTTTCAATTCGAGAGAGCTTAATGAATCAATGTCATT 4380  
DB 5354 GACGTGAAGTGAAGTGAATGATTTTCAATTCGAGAGAGCTTAATGAATCAATGTCATT 5413  
QY 4381 GCCAAAGTGTCAAAAGGAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTTACTAAATGGA 4440  
DB 5414 GCCAAAGTGTCAAAAGGAAGTTCTGATCTAGCAAAAGAAATATCAGGCTTTTACTAAATGGA 5473  
QY 4441 AAAAGTAAAAAATATTTTGAAGATGCTGTGTTTAAATACGATGGAATGTTTATTAATACG 4500  
DB 5474 AAAAGTAAAAAATATTTTGAAGATGCTGTGTTTAAATTAATACGATGGAATGTTTATTAATACG 5533  
QY 4501 AAAGAACAATTAAGAAAAAGCAAAAGAAAAAGAGAGCGGTCAATTGTAATGCTGCTTTA 4560

Db 5534 AAGGACAACTRGAAAAAGCAAAAGAAAAAGAGCGGCTATTGTAAATCTCTTTA 5593  
Qy 4561 TCGGTTGCTGAAACGGATTAATCCGCTGAGAGTAGCTATTGCAAGTCAATACTGTAAA 4620  
Db 5594 TCGGTTGCTGAAACGGATTAATCCGCTGAGAGTAGCTATTGCAAGTCAATACTGTAAA 5653  
Qy 4621 AATAAATTTAAAGCAGATTGATGAGAACCAATAAGAAAGCCGAGAGAGATTAATTCAT 4680  
Db 5654 AATAAATTTAAAGCAGATTGATGAGAACCAATAAGAAAGCCGAGAGAGATTAATTCAT 5713  
Qy 4681 GCGAAACATGTAAATGTGAGGCAAAATCATCTACTGTGTGTGTAATGCGCTCTGGA 4740  
Db 5714 GCGAAACATGTAAATGTGAGGCAAAATCATCTACTGTGTGTGTAATGCGCTCTGGA 5773  
Qy 4741 CTGCTATCAGCAAGATGCTTTTTCAGGAATGCGATCTGAGCATGCGAAGCTTATCA 4800  
Db 5774 CTGCTATCAGCAAGATGCTTTTTCAGGAATGCGATCTGAGCATGCGAAGCTTATCA 5833  
Qy 4801 AATGACAGATTGCAAAAGGTGATAAGGAAGATTTCTGCTGATTCCTTAAATGTGAAC 4860  
Db 5834 AATGACAGATTGCAAAAGGTGATAAGGAAGATTTCTGCTGATTCCTTAAATGTGAAC 5893  
Qy 4861 GGAATTAATTCATCTTGGGGTGAAATGTGGGGGAACCATGCGGCTCTCTTTCTAG 4920  
Db 5894 GGAATTAATTCATCTTGGGGTGAAATGTGGGGGAACCATGCGGCTCTCTTTCTAG 5953  
Qy 4921 GCGGTAGAGTGTCTTTGCGAATAATCTCTCATATATAAACTCTGCTTGTATTA 4980  
Db 5954 GCGGTAGAGTGTCTTTGCGAATAATCTCTCATATATAAACTCTGCTTGTATTA 6013  
Qy 4981 GGAACGAGGTAAATCTTTTAAAGTGAAGAAATCAAAAGTCAATGTACAGCTTGAAT 5040  
Db 6014 GGAACGAGGTAAATCTTTTAAAGTGAAGAAATCAAAAGTCAATGTACAGCTTGAAT 6073  
Qy 5041 GATTCATATTTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAAGAGCGCTGAATC 5100  
Db 6074 GATTCATATTTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAAGAGCGCTGAATC 6133  
Qy 5101 GGAGAAATGATCTGCTCAATCGTGTCTGATGAACGGAAGCTTTAGTATGATTC 5160  
Db 6134 GGAGAAATGATCTGCTCAATCGTGTCTGATGAACGGAAGCTTTAGTATGATTC 6193  
Qy 5161 GAGTTGAAGAGATTAATCTTTCAATGTATGATGCAAAAGATCAAAAAACAATTAATCA 5220  
Db 6194 GAGTTGAAGAGATTAATCTTTCAATGTATGCAAAAGATCAAAAAACAATTAATCA 6253  
Qy 5221 ATGCGCGGAATGCAAAATGAGAAAGCGCTGAGTTGAGCAACAGTTGCTCATACA 5280  
Db 6254 ATGCGCGGAATGCAAAATGAGAAAGCGCTGAGTTGAGCAACAGTTGCTCATACA 6313  
Qy 5281 AATATTGAAAAACAATCAGTTATAGCTATTGTAAAAAAACAGTAAATTAACAACGCGAAT 5340  
Db 6314 AATATTGAAAAACAATCAGTTATAGCTATTGTAAAAAAACAGTAAATTAACAACGCGAAT 6373  
Qy 5341 GATCAAGATGAAAAAATATCAATGTGATGCAAAAGATTATACTATGACCAATACTATA 5400  
Db 6374 GATCAAGATGAAAAAATATCAATGTGATGCAAAAGATTATACTATGACCAATACTATA 6433  
Qy 5401 GCAGTGCGAGTTGAGAGCAAAAGAGCCTCTGTGCAAGAGACTTTCGAATACTACC 5460  
Db 6434 GCAGTGCGAGTTGAGAGCAAAAGAGCCTCTGTGCAAGAGACTTTCGAATACTACC 6493  
Qy 5461 TTGAATTAAGACAGTTTCTTCTCATGTGATCAAACTGATATTGACAAAGATTTAGAGAA 5520  
Db 6494 TTGAATTAAGACAGTTTCTTCTCATGTGATCAAACTGATATTGACAAAGATTTAGAGAA 6553  
Qy 5521 GAAAAATATGAAATTAAGAAAAAGCAAAATGTTAATGTTTACGTGAAAAACAGAGTCAA 5580  
Db 6554 GAAAAATATGAAATTAAGAAAAAGCAAAATGTTAATGTTTACGTGAAAAACAGAGTCAA 6613  
Qy 5581 GTGTCACAAAATGCAAGTGTCTTCCGAGCAAGTGTGACAACTGTGAGATGAGAGCTGGA 5640

Db 6614 GTGTCACAAAATGCAAGTGTCTTCCGAGCAAGTGTGACAACTGTGAGATGAGAGCTGGA 6673  
Qy 5641 GTAGCAATTAATAAATTAACAATAATCTTCTGACATATATAAATAGTACTCAAAAT 5700  
Db 6674 GTAGCAATTAATAAATTAACAATAATCTTCTGACATATATAAATAGTACTCAAAAT 6733  
Qy 5701 GTACGAATGTCTTGGTAAAAAGCAAAATCTCATCTATTTAAACAATTTGAAATTGGA 5760  
Db 6734 GTACGAATGTCTTGGTAAAAAGCAAAATCTCATCTATTTAAACAATTTGAAATTGGA 6793  
Qy 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGTGAAGAGTTCGTAGCAATTAAGATTGTA 5820  
Db 6794 GCTGAGTTGAGAGCTGAGAGAGCTGAGTGAAGAGTTCGTAGCAATTAAGATTGTA 6853  
Qy 5821 AATAATTAAGATTAAGCAATTAATCAATGCAAAATCACTGCGCAAGGAAATGTGCGAGT 5880  
Db 6854 AATAATTAAGATTAAGCAATTAATCAATGCAAAATCACTGCGCAAGGAAATGTGCGAGT 6913  
Qy 5881 ATTACAGAGTCTGATGCGGTAATTCATATTAATGCAAGAACAGTGTGAGAGTGCCTGT 5940  
Db 6914 ATTACAGAGTCTGATGCGGTAATTCATATTAATGCAAGAACAGTGTGAGAGTGCCTGT 6973  
Qy 5941 GCAGCAATTAAGAGCCTCAACAGTGTGAATGAATTAACGATCTACAAACCATATGTA 6000  
Db 6974 GCAGCAATTAAGAGCCTCAACAGTGTGAATGAATTAACGATCTACAAACCATATGTA 7033  
Qy 6001 AAAGATTCTACAGTATTCCTTAAAGAAACAGATGATTAATTAATCTCAAGGCGAA 6060  
Db 7034 AAAGATTCTACAGTATTCCTTAAAGAAACAGATGATTAATTAATCTCAAGGCGAA 7093  
Qy 6061 GTAGATTAAGTGTGATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6120  
Db 7094 GTAGATTAAGTGTGATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7153  
Qy 6121 AAAAGAAAAATAGTAATTAATAAAGATTGTTACCAATAGTTACGTACTACATCTTTA 6180  
Db 7154 AAAAGAAAAATAGTAATTAATAAAGATTGTTACCAATAGTTACGTACTACATCTTTA 7213  
Qy 6181 AAATCTTTATTTGGCAATGCTCGCTGTTTCAAGCAACGCGAGTGCAGAAACTGTTAAT 6240  
Db 7214 AAATCTTTATTTGGCAATGCTCGCTGTTTCAAGCAACGCGAGTGCAGAAACTGTTAAT 7273  
Qy 6241 ATCAACAAGTTTATGAGAGAAACAGAGCTCTTGTGAAATCTTATTTAAATGCAAAA 6300  
Db 7274 ATCAACAAGTTTATGAGAGAAACAGAGCTCTTGTGAAATCTTATTTAAATGCAAAA 7333  
Qy 6301 CATTAATCTGTAAATCAGAGATTAACAGAAATTCATCGAGTAGTAGGTTCTGTGGT 6360  
Db 7334 CATTAATCTGTAAATCAGAGATTAACAGAAATTCATCGAGTAGTAGGTTCTGTGGT 7393  
Qy 6361 GTTGTGGAATGTAGAGTAGAGGCTTCTTGATACCAATTTATTAAGAAATATACC 6420  
Db 7394 GTTGTGGAATGTAGAGTAGAGGCTTCTTGATACCAATTTATTAAGAAATATACC 7453  
Qy 6421 AAGACAAGAGTTGGAATAAATCAATGTCGATGAAGGTTTGGAGAAGAAAGCTGAAATT 6480  
Db 7454 AAGACAAGAGTTGGAATAAATCAATGTCGATGAAGGTTTGGAGAAGAAAGCTGAAATT 7513  
Qy 6481 ACAGCAGATTTCTAAGCAAGAAATTTCTCTTTTGAAGTGTGAGTGCAGCAGCCGGGTA 6540  
Db 7514 ACAGCAGATTTCTAAGCAAGAAATTTCTCTTTTGAAGTGTGAGTGCAGCAGCCGGGTA 7573  
Qy 6541 GAGCCGAGTGTGAGAGAAACGTTTCCGTAAATCAATTTGCAAGAAAGACGAAATGTAT 6600  
Db 7574 GAGCCGAGTGTGAGAGAAACGTTTCCGTAAATCAATTTGCAAGAAAGACGAAATGTAT 7633  
Qy 6601 GTGGAAGAAAGCAAGATTTTGTGTAAGAAAGCTGAGATTACAGCAAAACGTTATAGTCT 6660  
Db 7634 GTGGAAGAAAGCAAGATTTTGTGTAAGAAAGCTGAGATTACAGCAAAACGTTATAGTCT 7693  
Qy 6661 GTTGCAATTTGGAATGCTCGAGTGTGAGAGTGTGCTCAAAAGAGAGTGTGAATTTGAGCAGA 6720  
Db 7694 GTTGCAATTTGGAATGCTCGAGTGTGAGAGTGTGCTCAAAAGAGAGTGTGAATTTGAGCAGA 7753

QY 6721 GTGGAGTTACCAAGATGATCAACAGAGACAGAGTGAATAATCTAAATTTATG 6780  
DB 7754 GTGGCAGTTACCAAGATGATCAACAGAGACAGAGTGAATAATCTAAATTTATG 7813  
QY 6781 ACTCGAACAAGTTAGATGTAATAGCAGAAAAAGAGATAAATCAAGTACTGGAATCGGT 6840  
DB 7814 ACTCGAACAAGTTAGATGTAATAGCAGAAAAAGAGATAAATCAAGTACTGGAATCGGT 7873  
QY 6841 TCAGCCGAGGCTGGAATCTTGACGCCGAGTATCTGAGTGGTCTTGCTCAATAATNT 6900  
DB 7874 TCAGCCGAGGCTGGAATCTTGACGCCGAGTATCTGAGTGGTCTTGCTCAATAATNT 7933  
QY 6901 GCAAAATTAAGTGAACAAGATATCGATCACTAGTACTTTACCTCTTCTAAGTAAAT 6960  
DB 7934 GCAAAATTAAGTGAACAAGATATCGATCACTAGTACTTTACCTCTTCTAAGTAAAT 7993  
QY 6961 GTAAAAAGCTCTTAATTAATAATTTTGAAATCTTGACACGCCGTGAGAGAGCCGACGCTT 7020  
DB 7994 GTAAAAAGCTCTTAATTAATAATTTTGAAATCTTGACACGCCGTGAGAGAGCCGACGCTT 8053  
QY 7021 GCAGAGTTACCGAGTGGTCTTGTTACACTATAATGTTCTGTATAGCTGAGTT 7080  
DB 8054 GCAGAGTTACCGAGTGGTCTTGTTACACTATAATGTTCTGTATAGCTGAGTT 8113  
QY 7081 CACAAATTAATCTGATTTGACTTCGGTACGAGAAAAAGTAAATGTAAACGCAAAAGAGAA 7140  
DB 8114 CACAAATTAATCTGATTTGACTTCGGTACGAGAAAAAGTAAATGTAAACGCAAAAGAGAA 8173  
QY 7141 AAAAAATTAATAGCAACAGACCAATGCGAATTCGAGAGACGCAATCGGACCAAT 7200  
DB 8174 AAAAAATTAATAGCAACAGACCAATGCGAATTCGAGAGACGCAATCGGACCAAT 8233  
QY 7201 GCTTGATTAATTAATTTTGAAACAGCTGAGAGAGTGAATAATTTGGAAGAAAGAA 7260  
DB 8234 GCTTGATTAATTAATTTTGAAACAGCTGAGAGAGTGAATAATTTGGAAGAAAGAA 8293  
QY 7261 ACAGAGTTTAAAAAATTTAGACAGAGTTAACAAGAAACAGATTAATAATGAT 7320  
DB 8294 ACAGAGTTTAAAAAATTTAGACAGAGTTAACAAGAAACAGATTAATAATGAT 8353  
QY 7321 GCTACGAAAAAATCTTACATCAGAGGTAATTTCTACAGAAATGATCTTGTAAGAGG 7380  
DB 8354 GCTACGAAAAAATCTTACATCAGAGGTAATTTCTACAGAAATGATCTTGTAAGAGG 8413  
QY 7381 GATAGAGGATATCTCAGGAGAGAGAAATTAAGCATGTTGTAAGATCTTGATTAAT 7440  
DB 8414 GATAGAGGATATCTCAGGAGAGAGAAATTAAGCATGTTGTAAGATCTTGATTAAT 8473  
QY 7441 GAAAAAATGTAGATATTTACACAGAGACAGAAATATATCATCTTACTGAGGTTTG 7500  
DB 8474 GAAAAAATGTAGATATTTACACAGAGACAGAAATATATCATCTTACTGAGGTTTG 8533  
QY 7501 GGAATCTGCAAGTCTTGCTCCGACATCAGAGACAGTGGCAGTTACAAATTTAAAGAAAT 7560  
DB 8534 GGAATCTGCAAGTCTTGCTCCGACATCAGAGACAGTGGCAGTTACAAATTTAAAGAAAT 8593  
QY 7561 TCCGAGTTACTGTTGAAAAATCTTTTGTGAAGACGCTGAAAAAGTAAATGTTAGATCG 7620  
DB 8594 TCCGAGTTACTGTTGAAAAATCTTTTGTGAAGACGCTGAAAAAGTAAATGTTAGATCG 8653  
QY 7621 GATATTACAGAAAAATGTTGCTTTAAACAGATATCAAGGCTCTGTAGAGACATTTGGAAAT 7680  
DB 8654 GATATTACAGAAAAATGTTGCTTTAAACAGATATCAAGGCTCTGTAGAGACATTTGGAAAT 8713  
QY 7681 GAGAGTGCCTATGACGAATTAATTTCTAATGGAAGATCAAAATATCATGTTAAAAATCT 7740  
DB 8714 GAGAGTGCCTATGACGAATTAATTTCTAATGGAAGATCAAAATATCATGTTAAAAATCT 8773  
QY 7741 AAGCTATTAGAAAAAATTTGATGTTATTTGTAAGAAATTAATCGGAATTTGAGACGGA 7800  
DB 8774 AAGCTATTAGAAAAAATTTGATGTTATTTGTAAGAAATTAATCGGAATTTGAGACGGA 8833

QY 7801 GCAAAAAGATTAAACCGTAGAGACGGTAGCTGCCGAGCCATTTATCTCAAAAGCAAGAAAT 7860  
DB 8834 GCAAAAAGATTAAACCGTAGAGACGGTAGCTGCCGAGCCATTTATCTCAAAAGCAAGAAAT 8893  
QY 7861 GAAATGAATTCAGAGGTTGAAAATGAGAGATATTTTCAATGAAGAAATAGATTAAT 7920  
DB 8894 GAAATGAATTCAGAGGTTGAAAATGAGAGATATTTTCAATGAAGAAATAGATTAAT 8953  
QY 7921 AGCCCTCTTAAAGGAATTTGGAAGAGAAATCAATGTCAAAGTGGAAAAAGAAAACAGAGTG 7980  
DB 8954 AGCCCTCTTAAAGGAATTTGGAAGAGAAATCAATGTCAAAGTGGAAAAAGAAAACAGAGTG 9013  
QY 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGTGAAGGCGCAGAAATTAATTTCCGAA 8040  
DB 9014 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGTGAAGGCGCAGAAATTAATTTCCGAA 9073  
QY 8041 GCAAAAAGTCCCGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100  
DB 9074 GCAAAAAGTCCCGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 9133  
QY 8101 CATGCAATTAATGTGAATATGGAAGCAACATTAATGAAAGTAAACAGCAGTTCTAAA 8160  
DB 9134 CATGCAATTAATGTGAATATGGAAGCAACATTAATGAAAGTAAACAGCAGTTCTAAA 9193  
QY 8161 GCAGTAAACAGGTTCTGTATTTGAGAGAGTGGAGTCAACAAAGCAGAACTTACTGCGA 8220  
DB 9194 GCAGTAAACAGGTTCTGTATTTGAGAGAGTGGAGTCAACAAAGCAGAACTTACTGCGA 9253  
QY 8221 GGTAAACTATGTGTAAGAGTTGAGAGAGAAATTTGTTCAAGAACAAATCGATTTGAATGCA 8280  
DB 9254 GGTAAACTATGTGTAAGAGTTGAGAGAGAAATTTGTTCAAGAACAAATCGATTTGAATGCA 9313  
QY 8281 ATTTCTAAAGTGAAGGTTTGGATGAATGAATGAATGAATGCTGTAATCTTGTAATATCA 8340  
DB 9314 ATTTCTAAAGTGAAGGTTTGGATGAATGAATGAATGAATGCTGTAATCTTGTAATATCA 9373  
QY 8341 GAAAAATGAGAGAGAAATTTCCGAGACAGAGTGAATTAATCTTACAGCACAAAAGTAATCT 8400  
DB 9374 GAAAAATGAGAGAGAAATTTCCGAGAGCAGAGTGAATTAATCTTACAGCACAAAAGTAATCT 9433  
QY 8401 GAATCCGTATGTTGTTTACGAAAGCAGATTAATGAATTAATGATTACACAAAAAATAT 8460  
DB 9434 GAATCCGTATGTTGTTTACGAAAGCAGATTAATGAATTAATGATTACACAAAAAATAT 9493  
QY 8461 ATTTGAGAGTCATGCTCTTGCTTAAATGATCAAAAGAAATGAAGCGAATTAAGATCT 8520  
DB 9494 ATTTGAGAGTCATGCTCTTGCTTAAATGATCAAAAGAAATGAAGCGAATTAAGATCT 9553  
QY 8521 TTAGCGTAGCCGAGTGTGCATGCAACAGAAACAAACAAAGATTTACAGAGATCAAAACAG 8580  
DB 9554 TTAGCGTAGCCGAGTGTGCATGCAACAGAAACAAACAAAGATTTACAGAGATCAAAACAG 9613  
QY 8581 TTAATCTTACACCTGTAAATGAGAGAAACGTATTTCAACTTCTGTCGCAAAAGCTTTGGCT 8640  
DB 9614 TTAATCTTACACCTGTAAATGAGAGAAACGTATTTCAACTTCTGTCGCAAAAGCTTTGGCT 9673  
QY 8641 AAAAAATGAATTTAGAAATGTAAAGAAAGTGAAGAGCTTGAAGCGGAGCA 8700  
DB 9674 AAAAAATGAATTTAGAAATGTAAAGAAAGTGAAGAGCTTGAAGCGGAGCA 9733  
QY 8701 GCAGCGTTGGAATTTATACAAAGAGTACTACAGAGCATTTGTTGACAGAAATTTGGAA 8760  
DB 9734 GCAGCGTTGGAATTTATACAAAGAGTACTACAGAGCATTTGTTGACAGAAATTTGGAA 9793  
QY 8761 ATTGAGATTAATTTAGAAACGATTTGCAAGAGATTAATGATTTGAAGTCAACGGAAC 8820  
DB 8821 ATTGAGATTAATTTAGAAACGATTTGCAAGAGATTAATGATTTGAAGTCAACGGAAC 8880  
QY 8821 GGAACCAAGAGAGCTTGTGCGAAAGAAATGATTTCTGTGGAATTAATTTCAAGG 8880  
DB 9854 GGAACCAAGAGAGCTTGTGCGAAAGAAATGATTTCTGTGGAATTAATTTCAAGG 9913  
QY 8881 GAAACAAATATCATCAATTGAAGATTAAGCAGAAATTTTGAACCGGAAGTGAATGTA 8940

```

Db 9914 GAAACAAATATCATTCATGAGATTAAGCCAGAAATTTGGAAACCGGAAAGGTAAATGTA 9973
Qy 8941 GATGCTTTGAAATGAATCTGATGTAGATCTACAGGAAAAAGTGGCTATGTTGGAAAT 9000
Db 9974 GATGCTTTGAAATGAATCTGATGTAGATCTACAGGAAAAAGTGGCTATGTTGGAAAT 10033
Qy 9001 GGTATTTGAAATGTTGATGTAAATATGATGATTAAGAAAAATGTGAAGCCAAATTCGGA 9060
Db 10034 GGTATTTGAAATGTTGATGTAAATATGATGATTAAGAAAAATGTGAAGCCAAATTCGGA 10093
Qy 9061 AGACATGCTATTTGAGAACTACCTGAAACCAAGATATCAAGCATTTACAGAGCAAAA 9120
Db 10094 AGACATGCTATTTGAGAACTACCTGAAACCAAGATATCAAGCATTTACAGAGCAAAA 10153
Qy 9121 GTAATATTTCTTGGAAAAAGAGACGCTGAGCTGAGCTGCAATATCGAATGTACACAT 9180
Db 10154 GTAATATTTCTTGGAAAAAGAGACGCTGAGCTGAGCTGCAATATCGAATGTACACAT 10213
Qy 9181 TCCAAATGAGATGATTTAAATAATTTGGCAAAAGCAGTATGATCTTTCTCAATTAATTAAC 9240
Db 10214 TCCAAATGAGATGATTTAAATAATTTGGCAAAAGCAGTATGATCTTTCTCAATTAATTAAC 10273
Qy 9241 AAAAATTCAAAAAATTAATTTTACCTTAGCATCAAGTATGATTCGAATGTGATTTGAT 9300
Db 10274 AAAAATTCAAAAAATTAATTTTACCTTAGCATCAAGTATGATTCGAATGTGATTTGAT 10333
Qy 9301 GGGGTGCTGAGAGCAAGAGTGCAGAGCCAAAGCGACAGTATGTAAAGATCAAAAT 9360
Db 10334 GGGGTGCTGAGAGCAAGAGTGCAGAGCCAAAGCGACAGTATGTAAAGATCAAAAT 10393
Qy 9361 AATAGAACTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAAAATCAATGTA 9420
Db 10394 AATAGAACTAATTAATGTTGATTTAGCAGAAAAATTTAAACAGAGGAAAAATCAATGTA 10453
Qy 9421 TATGCCGATATGATTAATAATTTAATTAATTAAGTAAAGCAAAATCTTAAGGCTTTGGGAT 9480
Db 10454 TATGCCGATATGATTAATAATTTAATTAATTAAGTAAAGCAAAATCTTAAGGCTTTGGGAT 10513
Qy 9481 GCCAAAGTCATGCTGACGCTCTTCGCAACTGCGCACTATTGAAAAAATGAAGTAA 9540
Db 10514 GCCAAAGTCATGCTGACGCTCTTCGCAACTGCGCACTATTGAAAAAATGAAGTAA 10573
Qy 9541 TTTAATTAATGCAATCCGAGAAATTTAAATAATTTGCAAGATTTGAGAGGAAAGCTAAT 9600
Db 10574 TTTAATTAATGCAATCCGAGAAATTTAAATAATTTGCAAGATTTGAGAGGAAAGCTAAT 10633
Qy 9601 AAAAAAAGCTCGGTAGAGATCTAATCAAGTATGATCTGATATACGATTAATATCATGCGAT 9660
Db 10634 AAAAAAAGCTCGGTAGAGATCTAATCAAGTATGATCTGATATACGATTAATATCATGCGAT 10693
Qy 9661 TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATCAAGAGAGAAAAAGGAAA 9720
Db 10694 TCTTCTGAAAAAGCATACAAAAAATTTGACATATCAATCAAGAGAGAAAAAGGAAA 10753
Qy 9721 AATAGA 9726
Db 10754 AATAGA 10759

```

```

RESULT 3
AA167642
ID AA167642 standard; DNA; 2780 BP.
XX
AC AA167642;
XX
DT 27-FEB-2002 (first entry)
XX
DE F. necrophorum truncated leukotoxin protein (SX) encoding DNA.
XX
KW Leukotoxin; infection; immunisation; liver abscess; foot rot;
KM bactericide; vaccine; truncated; SX; de.
XX

```

```

OS Fusobacterium necrophorum.
FH Key Location/Qualifiers
FT CDS 1..2780
FT /tag=a
FT /product="truncated leukotoxin (SX)"
XX
PN WO200180886-A2.
XX
PD 01-NOV-2001.
XX
PE 25-APR-2001; 2001WO-US013240.
XX
PR 25-APR-2000; 2000US-00558257.
PR 24-APR-2001; 2001US-00841786.
XX
PA (UNIV ) UNIV KANSAS STATE RES FOUND.
XX
PI Nagendra TG, Stewart GC, Narayanan SK, Chengappa MW;
XX
DR WPI: 2002-049245/06.
DR P-PSDB; AAG66007.
XX
PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an
PT animal against an infection e.g. foot rot, or liver abscesses caused by
PT the bacterium.
XX
PS Claim 10; Page 96-97; 108bp; English.
XX
CC The invention provides an isolated Fusobacterium necrophorum leukotoxin
CC polypeptide and its truncated versions. The leukotoxin protein can be
CC expressed by standard recombinant methodology. The leukotoxin gene is
CC useful for preparing a vaccine which confers effective immunity against
CC infection caused by F. necrophorum. The vaccine is useful for immunising
CC an animal against liver abscesses caused by F. necrophorum and for
CC preventing foot rot caused by F. necrophorum infection. The present
CC sequence represents a DNA encoding a F. necrophorum truncated leukotoxin
CC protein (SX)
XX
SQ Sequence 2780 BP; 1031 A; 362 C; 637 G; 750 T; 0 U; 0 Other;
XX
QX
XX
Query Match 28.6%; Score 2780; DB 6; Length 2780;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 919 GGGAGATTATTAAGAAAGAGGAGGAAAAAGAACTTATACACTCTTTAAGTTATCA 978
Db 1 GGGAGATTATTAAGAAAGAGGAGGAAAAAGAACTTATACACTCTTTAAGTTATCA 60
Qy 979 GATGTGAAGCTTCGTTAAGATTAATTAAGAAAAAGTATGAAAGAAATGTTGACATT 1038
Db 61 GATGTGAAGCTTCGTTAAGATTAATTAAGAAAAAGTATGAAAGAAATGTTGACATT 120
Qy 1039 AAGCTGAAGCAAGAAATTTCTATGATGCACTTTAGTACTAGCTTCAAGACACTCT 1098
Db 121 AAGCTGAAGCAAGAAATTTCTATGATGCACTTTAGTACTAGCTTCAAGACACTCT 180
Qy 1099 TTTAGCTTTGTACAGTCTTATTTCTCTATCAATTTAAATGATTTTAAAGTTATTG 1158
Db 181 TTTAGCTTTGTACAGTCTTATTTCTCTATCAATTTAAATGATTTTAAAGTTATTG 240
Qy 1159 ACAAGTACGACGATGTCCTTATTTGAAAAAGATCCCAAGTGAAGCAACAGAAAGAA 1218
Db 241 ACAAGTACGACGATGTCCTTATTTGAAAAAGATCCCAAGTGAAGCAACAGAAAGAA 300
Qy 1219 GCAATATTCATTTCTTACAGTGAAGTGAAGCACTATGAGAGCACTTCTCCATTA 1278
Db 301 GCAATATTCATTTCTTACAGTGAAGTGAAGCACTATGAGAGCACTTCTCCATTA 360
Qy 1279 AAAATTAACAATTTATTTATTTGAGAAAGCAATGAAACCTTCTCAGTATCGAGCGGA 1338
Db 361 AAAATTAACAATTTATTTATTTGAGAAAGCAATGAAACCTTCTCAGTATCGAGCGGA 420

```

QY 1339 TATATTTCTGCAAAAAGTAATTCATGTACTATTTGAAAGAAAGTAATAATCGAAGGGA 1398  
Db 421 TATATTTCTGCAAAAAGTAATTCATGTACTATTTGAAAGAAAGTAATAATCGAAGGGA 480  
QY 1339 AGAGAGATATTAATCTTCAAAATCTGAAAAATCTATTTGATGCTTCTGTTTCTGTTGGAAG 1458  
Db 481 AGAGAGATATTAATCTTCAAAATCTGAAAAATCTATTTGATGCTTCTGTTTCTGTTGGAAG 540  
QY 1459 ATGAGAGATTCATTAAGTAGCTCTTTTCAGTATTTGTTGACGGAAGGAAATTAATCT 1518  
Db 541 ATGAGAGATTCATTAAGTAGCTCTTTTCAGTATTTGTTGACGGAAGGAAATTAATCT 600  
QY 1519 TCCGTCAGATTTGCTTAAGGAGCAAAAAGTAGAATCGAAGACGATGATTAATGTGAGA 1578  
Db 601 TCCGTCAGATTTGCTTAAGGAGCAAAAAGTAGAATCGAAGACGATGATTAATGTGAGA 660  
QY 1579 AGTGAAGCATTAATTTCCATTTGAGCTGCTGTTAAAAAGGTGATTTGGGGGATAGTGTAT 1638  
Db 661 AGTGAAGCATTAATTTCCATTTGAGCTGCTGTTAAAAAGGTGATTTGGGGGATAGTGTAT 720  
QY 1639 GGGGTTGTGCTGCAAAATATTTCTAACTATATGCTTCTCCGATATAGATGTAGATGGA 1698  
Db 721 GGGGTTGTGCTGCAAAATATTTCTAACTATATGCTTCTCCGATATAGATGTAGATGGA 780  
QY 1639 TATCTACATGCCAAGAGCGACTAAATGTGAGGCTCATACATTACTAAAAATAGTGT 1758  
Db 781 TATCTACATGCCAAGAGCGACTAAATGTGAGGCTCATACATTACTAAAAATAGTGT 840  
QY 1759 CTGCAAAACGAGATCTGATTTGGGAACTTCCAGTTATAGATGATCAGTTTATGATCA 1818  
Db 841 CTGCAAAACGAGATCTGATTTGGGAACTTCCAGTTATAGATGATCAGTTTATGATCA 900  
QY 1819 GGTCACTCTAAAAATCAATTTTAGATGCATTAATAAACAAGCGTTTGGAGAGACAGTGTCA 1878  
Db 901 GGTCACTCTAAAAATCAATTTTAGATGCATTAATAAACAAGCGTTTGGAGAGACAGTGTCA 960  
QY 1879 GAGGAATTAAGAAATAGCTAACGAACCTTATTTAGTGTGCTGTCTGCAACCATAGCA 1938  
Db 961 GAGGAATTAAGAAATAGCTAACGAACCTTATTTAGTGTGCTGTCTGCAACCATAGCA 1020  
QY 1939 AATCATTAATTAATTTGCTTCTGTGCAATAGAGAGAGAGTGAAGACTTTCTTCAAGAGTG 1998  
Db 1021 AATCATTAATTAATTTGCTTCTGTGCAATAGAGAGAGAGTGAAGACTTTCTTCAAGAGTG 1080  
QY 1999 GAAAGGAGTAATGTAAAGGCACTTAATAGAGCTCAAAATCTTGCAGCGACTACGTCAAGT 2058  
Db 1081 GAAAGGAGTAATGTAAAGGCACTTAATAGAGCTCAAAATCTTGCAGCGACTACGTCAAGT 1140  
QY 2059 GGAAGTGTGCTGTACGAAGAGAGAAAAAGAACTTATTTGAAATGACAGAGTTTTT 2118  
Db 1141 GGAAGTGTGCTGTACGAAGAGAGAAAAAGAACTTATTTGAAATGACAGAGTTTTT 1200  
QY 2119 TATGGAATCTATAAAAATTAATGCTTGTGACAAATTTGCCGATCAGTGAATTTGGTATCG 2178  
Db 1201 TATGGAATCTATAAAAATTAATGCTTGTGACAAATTTGCCGATCAGTGAATTTGGTATCG 1260  
QY 2179 GAAAGGAAAAATGATATCAACAGTGAAAAATTAATATTAATAATCTTCAAAAAATG 2238  
Db 1261 GAAAGGAAAAATGATATCAACAGTGAAAAATTAATATTAATAATCTTCAAAAAATG 1320  
QY 2239 GCAAAAGCTGTTATGTATTAATTAAGAACTTTTAAAGAGAGCTTTTGAAGAGAAAAAGAA 2298  
Db 1321 GCAAAAGCTGTTATGTATTAATTAAGAACTTTTAAAGAGAGCTTTTGAAGAGAAAAAGAA 1380  
QY 2299 ACTCCAGAAATATGATCCGAAGATATTTGAATCTATTTGAAAAATTAATGATGATTTTCA 2358  
Db 1381 ACTCCAGAAATATGATCCGAAGATATTTGAATCTATTTGAAAAATTAATGATGATTTTCA 1440  
QY 2359 GAAAAATTTGATGAAAAACCGAGGCTTTTACTAAATGTGAAAGATGACATTAATTTCT 2418  
Db 1441 GAAAAATTTGATGAAAAACCGAGGCTTTTACTAAATGTGAAAGATGACATTAATTTCT 1500  
QY 2419 CCGGATGGAACCTTCAAAAAACAGGAACCTGCTATAGAAATTTGCAAACTATGTTGAGGAGAA 2478

Db 1501 CCGGATGGAACCTTCAAAAAACAGGAACCTGCTATAGAAATTTGCAAACTATGTTGAGGAGAA 1560  
QY 2479 ATGAAAAATTTAGAGAAAAATTTCCGAAAGAGATTTAAAGCTTTTTCAGAAAGATTGAGT 2538  
Db 1561 ATGAAAAATTTAGAGAAAAATTTCCGAAAGAGATTTAAAGCTTTTTCAGAAAGATTGAGT 1620  
QY 2539 GGAAGTATTAAGAAACCTTTGAATTTTACAGAGATAGGAAATTAAGCAATTTTCACT 2598  
Db 1621 GGAAGTATTAAGAAACCTTTGAATTTTACAGAGATAGGAAATTAAGCAATTTTCACT 1680  
QY 2599 TTTACCTTTTCCGAGCTTAATGAGAAAGAGATTTTCTCTGTGGAGAGAGCTGTTTCCG 2658  
Db 1681 TTTACCTTTTCCGAGCTTAATGAGAAAGAGATTTTCTCTGTGGAGAGAGCTGTTTCCG 1740  
QY 2659 TGGGTAGAACAGAGAAATTAAGCAAGGTATCCGTTGGAAAAAGGAGCTAACTTGCTGCA 2718  
Db 1741 TGGGTAGAACAGAGAAATTAAGCAAGGTATCCGTTGGAAAAAGGAGCTAACTTGCTGCA 1800  
QY 2719 AAAAAAGATTTAATTAATTAAGCTATCAATTAAGCAGAAACAGTGAATTTAGTGAAT 2778  
Db 1801 AAAAAAGATTTAATTAATTAAGCTATCAATTAAGCAGAAACAGTGAATTTAGTGAAT 1860  
QY 2779 ATTGGAATTTGCGAAGAGCATCATCCGAAAGTGCAGTCCGAGAGAGATTAATGTTCAA 2838  
Db 1861 ATTGGAATTTGCGAAGAGCATCATCCGAAAGTGCAGTCCGAGAGAGATTAATGTTCAA 1920  
QY 2839 AGATCGAAAAATTTGAGCTATGCTTGAAGCTAAAGAAAAAGCTGAATTTACAGAGAAAT 2898  
Db 1921 AGATCGAAAAATTTGAGCTATGCTTGAAGCTAAAGAAAAAGCTGAATTTACAGAGAAAT 1980  
QY 2899 ATTAATGAGATGATGAACAGACTTTTTCATGTAGCGGAGCTTTTAAATGTTGCTCA 2958  
Db 1981 ATTAATGAGATGATGAACAGACTTTTTCATGTAGCGGAGCTTTTAAATGTTGCTCA 2040  
QY 2959 GGTGGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGTATCACTTAAGCAAGATTT 3018  
Db 2041 GGTGGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGTATCACTTAAGCAAGATTT 2100  
QY 3019 TCCATGATGACGAAGCAATTTTGAAGCTTAATTAAGAAAAATTTGCTTTAAACAGTAAGAT 3078  
Db 2101 TCCATGATGACGAAGCAATTTTGAAGCTTAATTAAGAAAAATTTGCTTTAAACAGTAAGAT 2160  
QY 3079 GATACCTTCTGTTTGAAGATGCTGCGGTTCAAGCGGAAATCGGAACGAAAAATGCGGCGT 3138  
Db 2161 GATACCTTCTGTTTGAAGATGCTGCGGTTCAAGCGGAAATCGGAACGAAAAATGCGGCGT 2220  
QY 3139 GGGGTTGCTGTTGGGTAATGATTAATGATTAATTTTCAACAAAGCTTCCATGGAATTAAT 3198  
Db 2221 GGGGTTGCTGTTGGGTAATGATTAATGATTAATTTTCAACAAAGCTTCCATGGAATTAAT 2280  
QY 3199 GACGAGAGCAAGTAATTAATGATTAAGATTAAGATGTAAGTGAAGTAACTGCGGAA 3258  
Db 2281 GACGAGAGCAAGTAATTAATGATTAAGATTAAGATGTAAGTGAAGTAACTGCGGAA 2340  
QY 3259 TCTTTAGAGATGATGCAAAAAAGCAACGGAACAATCAACAGTATTTCTGTTCCGGAAGA 3318  
Db 2341 TCTTTAGAGATGATGCAAAAAAGCAACGGAACAATCAACAGTATTTCTGTTCCGGAAGA 2400  
QY 3319 ATTAATTAAGTTGGAAGTAAACCGAGTGAAGAAAAACGAAATCGAAGAAAAACAGAG 3378  
Db 2401 ATTAATTAAGTTGGAAGTAAACCGAGTGAAGAAAAACGAAATCGAAGAAAAACAGAG 2460  
QY 3379 GGAATTTTGGCAAAATCGGAAACAAAGTGAACCTGTGAAAAATTAATTAACGATAGT 3438  
Db 2461 GGAATTTTGGCAAAATCGGAAACAAAGTGAACCTGTGAAAAATTAATTAACGATAGT 2520  
QY 3439 ATGATTCATTAACGAAAAATTTACAAATTAATTTCTGAGAGAGTAAAAAAACCGGAG 3498  
Db 2521 ATGATTCATTAACGAAAAATTTACAAATTTACATTTCTGAAGAGGTAAAAAAACCGGAG 2580  
QY 3499 AATCTTCTTGAACGTTTCTCATACCTCCGATTAAGAACCGTCTTTCAGTTTGGAGCT 3558

Db 2581 AATCTTCTTGAAGCTTCTGACTCCGATAAAGACCGTCTTTGAGTGGAGCT 2640  
Qy 3559 TCTGGAAGTGTCTTTCAATATATATTAAGAAACATCTGCTGTGATGAGATG 3618  
Db 2641 TCTGGAAGTGTCTTTCAATATATATTAAGAAACATCTGCTGTGATGAGATG 2700  
Qy 3619 AAGATAATTTTGAAGGACCAATTAAGATGAGGTGACCTTCTTCAATCTTCTTT 3678  
Db 2701 AAGATAATTTTGAAGGACCAATTAAGATGAGGTGACCTTCTTCAATCTTCTTT 2760  
Qy 3679 GTTGAGCATGGGCGGATC 3698  
Db 2761 GTTGAGCATGGGCGGATC 2780

RESULT 4  
AA167645  
ID AA167645 standard; DNA; 2322 BP.  
XX  
AC AA167645;  
XX  
DT 27-FEB-2002 (first entry)  
XX  
DE F. necrophorum truncated leukotoxin protein (FINAL) encoding DNA.  
XX  
KW Leukotoxin; infection; immunisation; liver abscess; foot rot;  
KW bactericide; vaccine; truncated; FINAL; db.  
XX  
OS Fusbacterium necrophorum.  
XX  
FH Key Location/Qualifiers  
FT 1..2322  
FT CDS /product= "truncated leukotoxin (FINAL)"  
XX  
PN MO200180886-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 25-APR-2001; 2001MO-US013240.  
XX  
PR 25-APR-2000; 2000US-00558257.  
PR 24-APR-2001; 2001US-00841786.  
XX  
XX (UNIV ) UNIV KANSAS STATE RES FOUNDED.  
XX  
XX PA Negaraja TG, Stewart GC, Narayanan SK, Chengappa MM;  
XX  
XX PI WPI; 2002-049245/06.  
XX  
XX DR P-PSDB; AAG66010.  
XX  
XX PT Fusbacterium necrophorum polypeptide useful as vaccine in immunizing an  
XX  
XX PT animal against an infection e.g. foot rot, or liver abscesses caused by  
XX  
XX PT the bacterium.  
XX  
XX Claim 10; Page 100-101; 108bp; English.  
XX  
XX PS The invention provides an isolated Fusbacterium necrophorum leukotoxin  
XX  
XX CC polypeptide and its truncated versions. The leukotoxin protein can be  
XX  
XX CC expressed by standard recombinant methodology. The leukotoxin gene is  
XX  
XX CC useful for preparing a vaccine which confers effective immunity against  
XX  
XX CC infection caused by F. necrophorum. The vaccine is useful for immunising  
XX  
XX CC an animal against liver abscesses caused by F. necrophorum and for  
XX  
XX CC preventing foot rot caused by F. necrophorum infection. The present  
XX  
XX CC sequence represents a DNA encoding a F. necrophorum truncated leukotoxin  
XX  
XX CC protein (FINAL)  
XX  
SQ Sequence 2322 BP; 921 A; 299 C; 532 G; 570 T; 0 U; 0 Other;

Query Match 23.9%; Score 2322; DB 6; Length 2322;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7405 GGAATTTAAAGCATTGTGAAGACTTCTGATATTTATGGAAAAAATGATGATATTAACA 7464  
Db 1 GGAATTTAAAGCATTGTGAAGACTTCTGATATTTATGGAAAAAATGATGATATTAACA 60  
Qy 7465 GAGGACAAAGATTAATATACCTTCTAAGTGTGTTGGAACTGACGCTTCTTCCGA 7524  
Db 61 GAGGACAAAGATTAATATACCTTCTAAGTGTGTTGGAACTGACGCTTCTTCCGA 120  
Qy 7525 TCAGAACAGTGGCGAGTTACAAATTTTAAAGAAATTCGAGATTACGTTGAAATCT 7584  
Db 121 TCAGAACAGTGGCGAGTTACAAATTTTAAAGAAATTCGAGATTACGTTGAAATCT 180  
Qy 7585 TTTGTAAGACAGCTGAAAAAGTAAATGTTAGATCGATATTAACAGAAATGTTCTTTA 7644  
Db 181 TTTGTAAGACAGCTGAAAAAGTAAATGTTAGATCGATATTAACAGAAATGTTCTTTA 240  
Qy 7645 ACAGCATATCAAGATCTCTGAGACATTGGAAATAGAGCTGCTATGCAAAATTAAT 7704  
Db 241 ACAGCATATCAAGATCTCTGAGACATTGGAAATAGAGCTGCTATGCAAAATTAAT 300  
Qy 7705 TCTAATGGAAGATCAAAATATCAATTTTAAAGAAATTTAAGCTATTTAGAAATTTGAT 7764  
Db 301 TCTAATGGAAGATCAAAATATCAATTTTAAAGAAATTTAAGCTATTTAGAAATTTGAT 360  
Qy 7765 GTTATTTGTAAGATTAATTCGAAATTTGAGAGCGAAAGAAAGATTAACCGTAGAGCG 7824  
Db 361 GTTATTTGTAAGATTAATTCGAAATTTGAGAGCGAAAGAAAGATTAACCGTAGAGCG 420  
Qy 7825 GTAGTGCCTGGAGCCATTATCTCAAAAGCAAAAGATGAATTCAGAGTTGAAAT 7884  
Db 421 GTAGTGCCTGGAGCCATTATCTCAAAAGCAAAAGATGAATTCAGAGTTGAAAT 480  
Qy 7885 GAGAGAGATTTTCAATGAGAAATAGAGTACAGCCCTTCTTAAGAAATTTGAAAG 7944  
Db 481 GAGAGAGATTTTCAATGAGAAATAGAGTACAGCCCTTCTTAAGAAATTTGAAAG 540  
Qy 7945 GAAATCAATGTCAAGTGGAAAAAGAAACAGAGTACGCTGAATCTCAAGAGCTTCT 8004  
Db 541 GAAATCAATGTCAAGTGGAAAAAGAAACAGAGTACGCTGAATCTCAAGAGCTTCT 600  
Qy 8005 GTAGAGCAGTAGCAGAGGCGAGAAATTTATTCGAAAGCAAAAGATGCCGAACTTAT 8064  
Db 601 GTAGAGCAGTAGCAGAGGCGAGAAATTTATTCGAAAGCAAAAGATGCCGAACTTAT 660  
Qy 8065 TTGAAAGTTAGTACAAATCCGAAAGATTTTTCATGCGAGTAAATGTGAATATGAA 8124  
Db 661 TTGAAAGTTAGTACAAATCCGAAAGATTTTTCATGCGAGTAAATGTGAATATGAA 720  
Qy 8125 GCAACACATTAAGTGAAGTAAACAGAGTTCTAAAGCAGTAACAGGTTCTGATTTGGA 8184  
Db 721 GCAACACATTAAGTGAAGTAAACAGAGTTCTAAAGCAGTAACAGGTTCTGATTTGGA 780  
Qy 8185 GAGATTGAGTACCAAGGCGAGAGTACTGTCGAGTAAATCTATGTTAGAGTTGAG 8244  
Db 781 GAGATTGAGTACCAAGGCGAGAGTACTGTCGAGTAAATCTATGTTAGAGTTGAG 840  
Qy 8245 GAAAGAAATTTGTTCAAGAAATATCGATTTGAAATGCAATTTCTTAAGTGAAGTTGAT 8304  
Db 841 GAAAGAAATTTGTTCAAGAAATATCGATTTGAAATGCAATTTCTTAAGTGAAGTTGAT 900  
Qy 8305 GAAAGTAAAGTAACTGCTAAATCTTCTGATGATGAGAAATGAGAGGAAATTTGCCGA 8364  
Db 901 GAAAGTAAAGTAACTGCTAAATCTTCTGATGATGAGAAATGAGAGGAAATTTGCCGA 960  
Qy 8365 GCAGAGTAAATATCTTCTACAGACAAAGTAAATGCAATTCGTAATGCTTTACGAAG 8424  
Db 961 GCAGAGTAAATATCTTCTACAGACAAAGTAAATGCAATTCGTAATGCTTTACGAAG 1020  
Qy 8425 CAAAGTAAAGTAAATATGATTAACAAAGAAATATTTTCAAGATCAATGCTTTGCT 8484  
Db 1021 CAAAGTAAAGTAAATATGATTAACAAAGAAATATTTTCAAGATCAATGCTTTGCT 1080  
Qy 8485 TTAATGATTAACAAGAAATGAGGAAATATGAAATCTTTAGCGGTAGCCGATGTCATGCA 8544

```

Db      1081 TTAATGATACAAAGATGAAGCAATATAGAACTTTAGCGGTAGCCGGTGTGATGCA 1140
Qy      8545 CAAGGAACAAAGAAAGCATTTTCAGATTCGAACCAAGTTTACTTTCACACTGTAATGGA 8604
Db      1141 CAAGGAACAAAGAAAGCATTTTCAGATTCGAACCAAGTTTACTTTCACACTGTAATGGA 1200
Qy      8605 GGAAGAGTATCTCACTTGTGTCAGAAAGCTTGGCTTAAATGAAATTAATGGAATGTA 8664
Db      1201 GGAAGAGTATCTCACTTGTGTCAGAAAGCTTGGCTTAAATGAAATTAATGGAATGTA 1260
Qy      8665 AAAGGAAGTGAAGAGCTTTAGTTCGAGCGGAAACAGACGCCGTTGAAATTAACAAG 8724
Db      1261 AAAGGAAGTGAAGAGCTTTAGTTCGAGCGGAAACAGACGCCGTTGAAATTAACAAG 1320
Qy      8725 AGTACTACAGAGAGCATTTGGTTCAGAGAAATTGGGAAATTGGAGATTAATTAAGAAAGATT 8784
Db      1321 AGTACTACAGAGAGCATTTGGTTCAGAGAAATTGGGAAATTGGAGATTAATTAAGAAAGATT 1380
Qy      8785 GCAGAGATTAATACGATTTGTAAGAGTCAACGAGACGGAACCAAGAGAGGTCTTGTCCGA 8844
Db      1381 GCAGAGATTAATACGATTTGTAAGAGTCAACGAGACGGAACCAAGAGAGGTCTTGTCCGA 1440
Qy      8845 AAGATGGTATTTCTGTGAAAAATACATTTCAAGGGGAAACAAATTCATCCATTGAAGAT 8904
Db      1441 AAGATGGTATTTCTGTGAAAAATACATTTCAAGGGGAAACAAATTCATCCATTGAAGAT 1500
Qy      8905 AAAGCCAGAAATTGGTGAACCCGGAAGTGAATTTGTAAGAGCTTTGAATTAAGTAAGTA 8964
Db      1501 AAAGCCAGAAATTGGTGAACCCGGAAGTGAATTTGTAAGAGCTTTGAATTAAGTAAGTA 1560
Qy      8965 GATCTACAGAGAAAAAAGTGGTGGCTATGATGGAATTGGATTTGGAATTTGATGTAAT 9024
Db      1561 GATCTACAGAGAAAAAAGTGGTGGCTATGATGGAATTGGATTTGGAATTTGATGTAAT 1620
Qy      9025 AATGTGATTAAGAAAAATGTAGAACGCAAAATCGAAGACATGCTATTGTAGAAACTACT 9084
Db      1621 AATGTGATTAAGAAAAATGTAGAACGCAAAATCGAAGACATGCTATTGTAGAAACTACT 1680
Qy      9085 GGAAGAACAAAGATATCAAGCAATTTTCAAGAGCAAAAGTAAATTTCTGGAAAAAGAGAC 9144
Db      1681 GGAAGAACAAAGATATCAAGCAATTTTCAAGAGCAAAAGTAAATTTCTGGAAAAAGAGAC 1740
Qy      9145 GCTGACGTGACAGTGCATATCGAATGTACACATTTCCAAATGAGATGATTTAAAAAT 9204
Db      1741 GCTGACGTGACAGTGCATATCGAATGTACACATTTCCAAATGAGATGATTTAAAAAT 1800
Qy      9205 TTGGCAAGACAGTATGATCTTCTCAATTAATTAACCAAAATTTCAAAAAATTAATTAAT 9264
Db      1801 TTGGCAAGACAGTATGATCTTCTCAATTAATTAACCAAAATTTCAAAAAATTAATTAAT 1860
Qy      9265 TTAGCATCAAGTAGAATCGAATGTGCAATGTTCAATGGGGTGGCTGAAGCAAGAGTGA 9324
Db      1861 TTAGCATCAAGTAGAATCGAATGTGCAATGTTCAATGGGGTGGCTGAAGCAAGAGTGA 1920
Qy      9325 GGAAGCCAAAGCAGAGTATGATTAAGAAATCAAAATTAATGAAGTATTAATTTGATTTA 9384
Db      1921 GGAAGCCAAAGCAGAGTATGATTAAGAAATCAAAATTAATGAAGTATTAATTTGATTTA 1980
Qy      9385 GCAGGAAAAATTTAAACAGAGGAAACATCAATGTATATCGCGATATGATTAATAATTT 9444
Db      1981 GCAGGAAAAATTTAAACAGAGGAAACATCAATGTATATCGCGATATGATTAATAATTT 2040
Qy      9445 AATTAATTAAGTAAGCAAAATTTCTAAGGCTATTTGGGAGCCCAAAAGTCATGCTCAGTGT 9504
Db      2041 AATTAATTAAGTAAGCAAAATTTCTAAGGCTATTTGGGAGCCCAAAAGTCATGCTCAGTGT 2100
Qy      9505 TCAGCACTGACCACTATTGAAAAAATGAAGTAAATTTAATATGCGATCCGAGATTT 9564
Db      2101 TCAGCACTGACCACTATTGAAAAAATGAAGTAAATTTAATATGCGATCCGAGATTT 2160
Qy      9565 AAAAATTAATCTGCAAGATTGGAAGGAAAGCTAATTAATAAAAAAGTGGTATGATTAAT 9624

```

```

Db      2161 AAAATATATCTGCGAAGATTGGAAGGAAACCTATATAAAAAAGCTCGTAGATCTAAT 2220
Qy      9625 CAGGTAGACTGTGATTAAGGATTAATATATCATGGCATTTCTTGAAGAAAGCATACAAAAA 9684
Db      2221 CAGGTAGACTGTGATTAAGGATTAATATATCATGGCATTTCTTGAAGAAAGCATACAAAAA 2280
Qy      9685 TTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9726
Db      2281 TTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2322

RESULT 5
AA167643
ID   AA167643 standard; DNA; 2141 BP.
XX
XX
AC   AA167643;
XX
DT   27-FEB-2002 (first entry)
XX
DE   F. necrophorum truncated leukotoxin protein (GAS) DNA.
XX
KM   Leukotoxin; infection; immunisation; liver abscess; foot rot;
KM   bactericide; vaccine; truncated; GAS; de.
OS   Fusobacterium necrophorum.
PN   MO200180886-A2.
XX
PD   01-NOV-2001.
XX
PF   25-APR-2001; 2001WO-US013240.
XX
PR   25-APR-2000; 2000US-00558257.
XX
PR   24-APR-2001; 2001US-00841786.
XX
XX
PA   (UNIT) UNIV KANSAS STATE RES FOUND.
XX
PI   Nagaraja TG, Stewart GC, Narayanan SK, Chengappa MW;
XX
XX   WPI; 2002-049245/06.
XX
DR   Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an
PT   animal against an infection e.g. foot rot, or liver abscesses caused by
PT   the bacterium.
XX
XX
PS   Claim 10; Page 97-98; 108bp; English.
XX
CC   The invention provides an isolated Fusobacterium necrophorum leukotoxin
CC   polypeptide and its truncated versions. The leukotoxin protein can be
CC   expressed by standard recombinant methodology. The leukotoxin gene is
CC   useful for preparing a vaccine which confers effective immunity against
CC   infection caused by F. necrophorum. The vaccine is useful for immunising
CC   an animal against liver abscesses caused by F. necrophorum and for
CC   preventing foot rot caused by F. necrophorum infection. The present
CC   sequence represents a DNA encoding a F. necrophorum truncated leukotoxin
CC   protein (GAS)
XX
SQ   Sequence 2141 BP; 763 A; 310 C; 513 G; 555 T; 0 U; 0 Other;
Query Match      22.0%; Score 2141; DB 6; Length 2141;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      3553 GAGGCTTTCGAGAGGTTTCTTCAATATATTAATAAGAAACATCTGCTGTCTAGAT 3612
Db      1 GAGGCTTTCGAGAGGTTTCTTCAATATATTAATAAGAAACATCTGCTGTCTAGAT 60
Qy      3613 GAGATTAAGATTAATTTGAAGAGAGCAATTAATAAGTGAAGTGAAGTCTTCTGATTTCT 3672
Db      61 GAGATTAAGATTAATTTGAAGAGAGCAATTAATAAGTGAAGTGAAGTCTTCTGATTTCT 120
Qy      3673 ACTTTGTTGAGAGATGAGGCGGATCTGCACTTCAGTGGAAATCAATTTGAAAGTGA 3732

```

Db 121 ACTTTGTGAGCATGGGGGCGGATCTGTCACCTTCAGTGGAATCATATTGGAAGTGA 180  
Qy 3733 AATAGCAACATCAGTGTGTTTACGTGAGCGGCTGCTGTAATTAATTCAAAAGTAA 3792  
Db 181 AATAGCAACATCAGTGTGTTTACGTGAGCGGCTGCTGTAATTAATTCAAAAGTAA 240  
Qy 3793 ACAATGCTTTGGTTTAAAAATAGTATTAATGAAATGCCAATTAATTAAGTAAATGCT 3852  
Db 241 ACAATGCTTTGGTTTAAAAATAGTATTAATGAAATGCCAATTAATTAAGTAAATGCT 300  
Qy 3853 TTGAGTGGAGAACTCAAGTAGCAGCAGAGAGGTTTGGAAAGCAGTTAAAGAAAGTGA 3912  
Db 301 TTGAGTGGAGAACTCAAGTAGCAGCAGAGAGGTTTGGAAAGCAGTTAAAGAAAGTGA 360  
Qy 3913 GGACAGAGAAAAAGTTATCTATTGGGAACTTGCTTATCACTTACTAGTGAACATGAA 3972  
Db 361 GGACAGAGAAAAAGTTATCTATTGGGAACTTGCTTATCACTTACTAGTGAACATGAA 420  
Qy 3973 GTTTCGCAAAATCGAAAAATTAACAGTAGAGAGAAATCGAAAGCCAAAAATGAT 4032  
Db 421 GTTTCGCAAAATCGAAAAATTAACAGTAGAGAGAAATCGAAAGCCAAAAATGAT 480  
Qy 4033 GTTGAATGCTGCTTATCAAGCGGACACCAAGTAGCAGAGCCTTTAAATTACAGCT 4092  
Db 481 GTTGAATGCTGCTTATCAAGCGGACACCAAGTAGCAGAGCCTTTAAATTACAGCT 540  
Qy 4093 GGAAGTCAAAATGGAATGTAAGGAGCTAGCTGTAAGTCCAAATTAACAACAAGTA 4152  
Db 541 GGAAGTCAAAATGGAATGTAAGGAGCTAGCTGTAAGTCCAAATTAACAACAAGTA 600  
Qy 4153 AATGCTTCTATTAGTGTGGAGATTAATCAACGTTAATCGACCGAACCAAGCTCTT 4212  
Db 601 AATGCTTCTATTAGTGTGGAGATTAATCAACGTTAATCGACCGAACCAAGCTCTT 660  
Qy 4213 TTAGCAACCACTCAAGTGACTGTCAGTAGAGAGCGGAGGCAATTAGTCTGGAGG 4272  
Db 661 TTAGCAACCACTCAAGTGACTGTCAGTAGAGAGCGGAGGCAATTAGTCTGGAGG 720  
Qy 4273 GGATTAGAAATTAATCAAGGAGCTGTTTGTCAATTAAGTGAACAATGACGTGAAGCT 4332  
Db 721 GGATTAGAAATTAATCAAGGAGCTGTTTGTCAATTAAGTGAACAATGACGTGAAGCT 780  
Qy 4333 AGCGTTGATTAATCTTCATCGAAGAGCTAATGAATCAATGTCATTGCCAAAGTATGTC 4392  
Db 781 AGCGTTGATTAATCTTCATCGAAGAGCTAATGAATCAATGTCATTGCCAAAGTATGTC 840  
Qy 4393 AAAGGAAGTTCGATGATCAAAAGAAATACAGGCTTACTAATGGAAGATTAAGAA 4452  
Db 841 AAAGGAAGTTCGATGATCAAAAGAAATACAGGCTTACTAATGGAAGATTAAGAA 900  
Qy 4453 TATTTAGAGATCGTGTATTAATACGACTGGAATGTTATTTATCGAAGGAACAATCA 4512  
Db 901 TATTTAGAGATCGTGTATTAATACGACTGGAATGTTATTTATCGAAGGAACAATCA 960  
Qy 4513 GAAAAAGCAAAAGAAAAAGAGAGCGGTCATTTGTAATGTCGCTTTATCGGTTGCTGA 4572  
Db 961 GAAAAAGCAAAAGAAAAAGAGAGCGGTCATTTGTAATGTCGCTTTATCGGTTGCTGA 1020  
Qy 4573 ACCGATTAATCCGCTGGAAGAGATGCTAATGCGAATTAATCTGTTAAAAATTAATTA 4632  
Db 1021 ACCGATTAATCCGCTGGAAGAGATGCTAATGCGAATTAATCTGTTAAAAATTAATTA 1080  
Qy 4633 GCAGATTGAGTGAAGCAATAAGAAAGCGAGAGATTAATTAATCATCGAAGCAATGTA 4692  
Db 1081 GCAGATTGAGTGAAGCAATAAGAAAGCGAGAGATTAATTAATCATCGAAGCAATGTA 1140  
Qy 4693 AATGTGAGGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCTGACTTATCAGC 4752  
Db 1141 AATGTGAGGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCTGACTTATCAGC 1200  
Qy 4753 AAAGATGCTTTTCAAGAAATGGAGTCTGAGCATGGAAGACTTATCAAAATGACAGATT 4812  
Db 1201 AAAGATGCTTTTCAAGAAATGGAGTCTGAGCATGGAAGACTTATCAAAATGACAGATT 1260

Qy 4813 GCAAGGTGATTAAGGAAGAAATTTCTGCTGATTCTTAAATGGAAGCAAAATTAATCC 4872  
Db 1261 GCAAGGTGATTAAGGAAGAAATTTCTGCTGATTCTTAAATGGAAGCAAAATTAATCC 1320  
Qy 4873 ATTCTTGGGGTGAATGTTGCGGGAACCAATGCCGTTCTCTTCTTACGCGGTAGAGCT 4932  
Db 1321 ATTCTTGGGGTGAATGTTGCGGGAACCAATGCCGTTCTCTTCTTACGCGGTAGAGCT 1380  
Qy 4933 GCTTTGCGAATTAATCTCTTCAATTAATAACCTCTGCTTTGATTACGGAACGAAGTA 4992  
Db 1381 GCTTTGCGAATTAATCTCTTCAATTAATAACCTCTGCTTTGATTACGGAACGAAGTA 1440  
Qy 4993 AATCCTTTTAACTGGAAGAAATTAACAAGTCAATGTAACAAGCTTTGAATGATTCTCATATT 5052  
Db 1441 AATCCTTTTAACTGGAAGAAATTAACAAGTCAATGTAACAAGCTTTGAATGATTCTCATATT 1500  
Qy 5053 ACAAAGTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGCTGGAATCGAGAAATGTA 5112  
Db 1501 ACAAAGTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGCTGGAATCGAGAAATGTA 1560  
Qy 5113 TCTGCAATCGTGTGTTCTGATGAACGGAAGCTTTAGTTAGTATTCTGAGTTGAAGTA 5172  
Db 1561 TCTGCAATCGTGTGTTCTGATGAACGGAAGCTTTAGTTAGTATTCTGAGTTGAAGTA 1620  
Qy 5173 GTAAGTCTTTCAATGTAATGCAAAAGATCAAAATAACAATTAATTAATGCGGGAAT 5232  
Db 1621 GTAAGTCTTTCAATGTAATGCAAAAGATCAAAATAACAATTAATTAATGCGGGAAT 1680  
Qy 5233 GCAAAATGAGAAAAAGCGGCTGAGTGTGAGCAACAGTTGCTCATACAAATTTGGAANA 5292  
Db 1681 GCAAAATGAGAAAAAGCGGCTGAGTGTGAGCAACAGTTGCTCATACAAATTTGGAANA 1740  
Qy 5293 CAATCAGTTATAGCTATTGTAATAAAACGTAATAATTAACAACGCGAATGATCAAGTGA 5352  
Db 1741 CAATCAGTTATAGCTATTGTAATAAAACGTAATAATTAACAACGCGAATGATCAAGTGA 1800  
Qy 5353 AAAAATATCAATGTACATGCAAAAGATTAATTAATGACCAATTAATGACAGTCCGAGTT 5412  
Db 1801 AAAAATATCAATGTACATGCAAAAGATTAATTAATGACCAATTAATGACAGTCCGAGTT 1860  
Qy 5413 GGAGAGCAAAAGAGCGCTGTCGCAAGAGCTTCTGCAAGTACTTGAATTAAGACA 5472  
Db 1861 GGAGAGCAAAAGAGCGCTGTCGCAAGAGCTTCTGCAAGTACTTGAATTAAGACA 1920  
Qy 5473 GTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGGAAGAAATTAATGA 5532  
Db 1921 GTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGGAAGAAATTAATGA 1980  
Qy 5533 AATTAAGAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACAGATCAAGTGTCACAAT 5592  
Db 1981 AATTAAGAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACAGATCAAGTGTCACAAT 2040  
Qy 5593 GCGAGAGCTTCCGAGCAAGTGAAGCAAGCTGAGTGAAGCTGAGTGAAGTAAAT 5652  
Db 2041 GCGAGAGCTTCCGAGCAAGTGAAGCAAGCTGAGTGAAGCTGAGTGAAGTAAAT 2100  
Qy 5653 AAAATTACAAAATTAATCTGCAATATTAATAATTAATAGTAC 5693  
Db 2101 AAAATTACAAAATTAATCTGCAATATTAATAATTAATAGTAC 2141

RESULT 6  
AA167644  
ID AA167644 standard; DNA; 1887 BP.  
XX  
AC AA167644;  
XX  
DT 27-FEB-2002 (first entry)  
XX  
DE F. necrophorum truncated leukotoxin protein (SH) encoding DNA.  
XX  
KW Leukotoxin; infection; immunisation; liver abscess; foot rot;

KM bactericide; vaccine; truncated; SH; ds.  
XX  
OS Fusobacterium necrophorum.  
XX  
FH Key Location/Qualifiers  
FT CDS 3..1886  
FT /tag=a  
FT /product="truncated leukotoxin (SH)"  
PN WO200180886-A2.  
XX  
PD 01-NOV-2001.  
XX  
PF 25-APR-2001; 2001WO-US013240.  
XX  
PR 25-APR-2000; 2000US-00558257.  
XX  
PR 24-APR-2001; 2001US-00841786.  
XX  
PA (UNITV ) UNIV KANSAS STATE RES FOUND.  
XX  
PI Nagaraja TG, Stewart GC, Narayanan SK, Chengappa MM;  
XX  
DR MPI: 2002-049245/06.  
XX  
DR P-PSDB; AAG66009.  
XX  
PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an  
XX animal against an infection e.g. foot rot, or liver abscesses caused by  
XX the bacterium.  
PS Claim 10; Page 98-99; 108bp; English.  
XX  
CC The invention provides an isolated Fusobacterium necrophorum leukotoxin  
CC polypeptide and its truncated versions. The leukotoxin protein can be  
CC expressed by standard recombinant methodology. The leukotoxin gene is  
CC useful for preparing a vaccine which confers effective immunity against  
CC infection caused by F. necrophorum. The vaccine is useful for immunizing  
CC an animal against liver abscesses caused by F. necrophorum and for  
CC preventing foot rot caused by F. necrophorum infection. The present  
CC sequence represents a DNA encoding a F. necrophorum truncated leukotoxin  
CC protein (SH)  
XX  
SQ Sequence 1887 BP; 712 A; 268 C; 439 G; 468 T; 0 U; 0 Other;  
Query Match 19.4%; Score 1887; DB 6; Length 1887;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1887; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5624 CTGCAAGTAGAGCTGGAGTGAAGCAATTAATTAACAAATTAATCTTCTGCAATATTA 5683  
DB 1 CTGCAAGTAGAGCTGGAGTGAAGCAATTAATTAACAAATTAATCTTCTGCAATATTA 60  
QY 5684 AAAATAGTACTCAAAATGTAAGCAATGCTTGGTAAAGCAAAATCTCATCTATTA 5743  
DB 61 AAAATAGTACTCAAAATGTAAGCAATGCTTGGTAAAGCAAAATCTCATCTATTA 120  
QY 5744 AAACAATTGGAATTGGAGCTGAGTGAAGCTGAGAGAGCTGAGTGAAGAGTTCTGTAG 5803  
DB 121 AAACAATTGGAATTGGAGCTGAGTGAAGCTGAGAGAGCTGAGTGAAGAGTTCTGTAG 180  
QY 5804 CAGTGAAATGAATTTGAATTAATAGATAGCAAGATTAATCATGCAAAATCACTGCGA 5863  
DB 181 CAGTGAAATGAATTTGAATTAATAGATAGCAAGATTAATCATGCAAAATCACTGCGA 240  
QY 5864 AAGGAATGCGAGTATTACAGAGTCTGATGCGGTATTTGCTAATTAAGAGAAACG 5923  
DB 241 AAGGAATGCGAGTATTACAGAGTCTGATGCGGTATTTGCTAATTAAGAGAAACG 300  
QY 5924 TGCTGAGAGTGGCCCGTGCAGCAATAGAGCTCAACAGTGTGATGAATTAAGAGAT 5983  
DB 301 TGCTGAGAGTGGCCCGTGCAGCAATAGAGCTCAACAGTGTGATGAATTAAGAGAT 360  
QY 5984 CTACAAAGCATATGTAAGAAAGATTCTACAGTATTTGCTAAAGAGAAACAGATGATTATA 6043

DB 361 CTACAAAGCATATGTAAGAAAGATTCTACAGTATTTGCTAAAGAGAAACAGATGATTATA 420  
QY 6044 TTAAGTCTCAAGGCGCAAGTAGATTAAGAGTGAATTAAGTATTCATAATTAATTA 6103  
DB 421 TTAAGTCTCAAGGCGCAAGTAGATTAAGAGTGAATTAAGTATTCATAATTAATTA 480  
QY 6104 ACGAAGCTTATCACAAAAAAGAAAAATTAAGTATTAAGAAAAAGTTGTTACCAATAGTT 6163  
DB 481 ACGAAGCTTATCACAAAAAAGAAAAATTAAGTATTAAGAAAAAGTTGTTACCAATAGTT 540  
QY 6164 CAGCTACTCATATCTTTAAATCTTTATTTGCGAAATGCGCGTGTTCAGAGCAAGCCGAG 6223  
DB 541 CAGCTACTCATATCTTTAAATCTTTATTTGCGAAATGCGCGTGTTCAGAGCAAGCCGAG 600  
QY 6224 TGGCAGAACTGTTAATATCAACAAGGTTTATGAGAAACAGAAAGCTTTGTAGAAATTT 6283  
DB 601 TGGCAGAACTGTTAATATCAACAAGGTTTATGAGAAACAGAAAGCTTTGTAGAAATTT 660  
QY 6284 CTATATTTAATGCAAAACATTATCTGTAAATCAGAGATTAACAAGATTCATTCGAG 6343  
DB 661 CTATATTTAATGCAAAACATTATCTGTAAATCAGAGATTAACAAGATTCATTCGAG 720  
QY 6344 TAGTAGGTTCTGTGTGTGTGTGTGTGAATAATGTAAGAGTGAAGCTTTCTGTATACATA 6403  
DB 721 TAGTAGGTTCTGTGTGTGTGTGTGTGTGAATAATGTAAGAGTGAAGCTTTCTGTATACATA 780  
QY 6404 TTATTAATAAAGAAATACCAAGCAAGAGTGAAGAAATCTACATGTCTGATGAAGTTTCG 6463  
DB 781 TTATTAATAAAGAAATACCAAGCAAGAGTGAAGAAATCTACATGTCTGATGAAGTTTCG 840  
QY 6464 GAGAAAGCTGAATTAACAGAGATTTCTAAGCAAGAAATTTCTTTTGAAGTGGAG 6523  
DB 841 GAGAAAGCTGAATTAACAGAGATTTCTAAGCAAGAAATTTCTTTTGAAGTGGAG 900  
QY 6524 TCGCAGCAGCCGGGTGAGACCCGAGTGAAGAGAACCGTTTCCGTAAATGAATTTGAG 6583  
DB 901 TCGCAGCAGCCGGGTGAGACCCGAGTGAAGAGAACCGTTTCCGTAAATGAATTTGAG 960  
QY 6584 GAAAGACGGAAGTATGATGGAAGAGCAAGATTTTGTATTAAGAAAAAGCTGAGATTACAG 6643  
DB 961 GAAAGACGGAAGTATGATGGAAGAGCAAGATTTTGTATTAAGAAAAAGCTGAGATTACAG 1020  
QY 6644 CAAAACGTTATAGTTCTGTGTGCAATTTGCAATGCGCAGTGGAGTGGCTGCAAAAGAG 6703  
DB 1021 CAAAACGTTATAGTTCTGTGTGCAATTTGCAATGCGCAGTGGAGTGGCTGCAAAAGAG 1080  
QY 6704 CTGGAATTGAGCAGCAGTGGCAGTATCCAAAGATGAATCAAAACGAGACCAAGAGTGA 6763  
DB 1081 CTGGAATTGAGCAGCAGTGGCAGTATCCAAAGATGAATCAAAACGAGACCAAGAGTGA 1140  
QY 6764 AAAATTCTAAATTAATGACTGCAAAACAAGTATGATGTAATGAGCAAAATGAGATTAAT 6823  
DB 1141 AAAATTCTAAATTAATGACTGCAAAACAAGTATGATGTAATGAGCAAAATGAGATTAAT 1200  
QY 6824 CAGGTACTGGAATCGGTTCAAGCCGAGAGCTGGAATTTTTCAGCCGAGATTCGAGTGG 6883  
DB 1201 CAGGTACTGGAATCGGTTCAAGCCGAGAGCTGGAATTTTTCAGCCGAGATTCGAGTGG 1260  
QY 6884 TTTCTGTAAATTAATTTGCAAAATTAAGTGAAGAAACAGATATGATATAGTACTTAACT 6943  
DB 1261 TTTCTGTAAATTAATTTGCAAAATTAAGTGAAGAAACAGATATGATATAGTACTTAACT 1320  
QY 6944 CTTCTACTGATGTAATTAATGTAAGAGCTCTTAATTAATTTTGAATTCCTTGAACGCGGTG 7003  
DB 1321 CTTCTACTGATGTAATTAATGTAAGAGCTCTTAATTAATTTTGAATTCCTTGAACGCGGTG 1380  
QY 7004 GAGAGCGCAGAGTCTTTCAGCAGAGTTCAGAGTGGTTTCTGTAAACATTAATAGTT 7063  
DB 1381 GAGAGCGCAGAGTCTTTCAGCAGAGTTCAGAGTGGTTTCTGTAAACATTAATAGTT 1440  
QY 7064 CTGTGATAGCTGAGTTTACAAATTAATCTGATTTTGAATTTCCGTACGAGAAAAAGTAAATG 7123  
DB 1441 CTGTGATAGCTGAGTTTACAAATTAATCTGATTTTGAATTTCCGTACGAGAAAAAGTAAATG 1500

QY 7124 TAACGGCAAAAGAGAAAAATATTAGCAAAACAGCAAAATGCAAGATCGAGGAG 7183  
DB 1501 TAACGGCAAAAGAGAAAAATATTAGCAAAACAGCAAAATGCAAGATCGAGGAG 1560  
QY 7184 CAGCAATCGAGACCAATGCTTGGTAAATATTTTGAACAGCTGTAGAGATGAAAAA 7243  
DB 1561 CAGCAATCGAGACCAATGCTTGGTAAATATTTTGAACAGCTGTAGAGATGAAAAA 1620  
QY 7244 ATTCTGAAGAAAGAGAAAGAAAGTTTAAACCTTTAGACGAAGTTAACAAAGAACAG 7303  
DB 1621 ATTCTGAAGAAAGAGAAAGAAAGTTTAAACCTTTAGACGAAGTTAACAAAGAACAG 1680  
QY 7304 ATAAAAAGTAATATGATGTACGAAAAAAATCTTACAACTACAGATATTTTACAGAG 7363  
DB 1681 ATAAAAAGTAATATGATGTACGAAAAAAATCTTACAACTACAGATATTTTACAGAG 1740  
QY 7364 ATACTTCTGTAAAGGAGATAGAGAGATACAGGAGAGAGAAATTAAGCCATTGTGA 7423  
DB 1741 ATACTTCTGTAAAGGAGATAGAGAGATACAGGAGAGAGAAATTAAGCCATTGTGA 1800  
QY 7424 AGACTTCTGTATATTATTTGAAAAAATGTAGATATTACAAAGAGAAAGAAATATATCA 7483  
DB 1801 AGACTTCTGTATATTATTTGAAAAAATGTAGATATTACAAAGAGAAAGAAATATATCA 1860  
QY 7484 CTTCTACTGTGTGTTGGGAACTGCAG 7510  
DB 1861 CTTCTACTGTGTGTTGGGAACTGCAG 1887

## RESULT 7

AA167641  
ID AA167641 standard; DNA; 1130 BP.

AA167641;

27-FEB-2002 (first entry)

F. necrophorum truncated leukotoxin protein (BSSE) encoding DNA.

Leukotoxin; infection; immunisation; liver abscess; foot rot;

bactericide; vaccine; truncated; BSSE; db.

Fusobacterium necrophorum.

Key Location/Qualifiers

FT CDS 1..1110

FT /tag= a

FT /product= "truncated leukotoxin (BSSE)"

PN WO200180886-A2.

PD 01-NOV-2001.

PF 25-APR-2001; 2001MO-US013240.

PR 25-APR-2000; 2000US-00558257.

PR 24-APR-2001; 2001US-00841786.

XX (UNIV ) UNIV KANSAS STATE RES FOUNDED.

XX Nagaraia TG, Stewart GC, Narayanan SK, Chengappa MM;

XX WPI; 2002-049245/06.

XX P-PSDB; AAG66006.

PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing an

PT animal against an infection e.g. foot rot, or liver abscesses caused by

PT the bacterium.

XX Claim 10; Page 95; 108pp; English.

XX The invention provides an isolated Fusobacterium necrophorum leukotoxin

CC polypeptide and its truncated versions. The leukotoxin protein can be  
CC expressed by standard recombinant methodology. The leukotoxin gene is  
CC useful for preparing a vaccine which confers effective immunity against  
CC infection caused by F. necrophorum. The vaccine is useful for immunising  
CC an animal against liver abscesses caused by F. necrophorum and for  
CC preventing foot rot caused by F. necrophorum infection. The present  
CC sequence represents a DNA encoding a F. necrophorum truncated leukotoxin  
CC protein (BSSE)

XX SQ Sequence 1130 BP; 444 A; 135 C; 234 G; 317 T; 0 U; 0 Other;

Query Match 11.6%; Score 1130; DB 6; Length 1130;

Best Local Similarity 100.0%; Pred. No. 9.1e-210;

Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGGCATCAAAATTAAGTTAGAGACAAAGAAAGATTCAGATTCTAAAAA 60  
DB 1 ATGACGGCATCAAAATTAAGTTAGAGACAAAGAAAGATTCAGATTCTAAAAA 60  
QY 61 GTTTTAATGATTTGGGATGTTGATTAACCTATGACGGGTATGATCAATC 120  
DB 61 GTTTTAATGATTTGGGATGTTGATTAACCTATGACGGGTATGATCAATC 120  
QY 121 ACCGGAGTGAATTTTGGAAACAAAAATAGAAAAAGATATGTTATGACATTACT 180  
DB 121 ACCGGAGTGAATTTTGGAAACAAAAATAGAAAAAGATATGTTATGACATTACT 180  
QY 181 ACAACCAAGATTCAAGGGGAGAACCTTTTAACTATTAATAGTTGCTTTACAGAA 240  
DB 181 ACAACCAAGATTCAAGGGGAGAACCTTTTAACTATTAATAGTTGCTTTACAGAA 240  
QY 241 AATAATATAGCAAAATCTATATTGTTGGGAAAAAGATAGTACGGGGTAAATATCTTTT 300  
DB 241 AATAATATAGCAAAATCTATATTGTTGGGAAAAAGATAGTACGGGGTAAATATCTTTT 300  
QY 301 AACTTGTCAATGAAAAAATTAAGTAGATGAGATTAATCAAGAAATTCAGAAAAATATA 360  
DB 301 AACTTGTCAATGAAAAAATTAAGTAGATGAGATTAATCAAGAAATTCAGAAAAATATA 360  
QY 361 ATTGAGAGAAATTTATATTTCTTAACTCTGGAAGGATGCGATGAGAAAAATGAGATT 420  
DB 361 ATTGAGAGAAATTTATATTTCTTAACTCTGGAAGGATGCGATGAGAAAAATGAGATT 420  
QY 421 ATCAATGCTGCTTTCTTCAATCTATTTTCCAAAAAGATGATTTAAGAGGCTTGG 480  
DB 421 ATCAATGCTGCTTTCTTCAATCTATTTTCCAAAAAGATGATTTAAGAGGCTTGG 480  
QY 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGAAATCAATCTGTTGAAGC 540  
DB 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGAAATCAATCTGTTGAAGC 540  
QY 541 ATTCATTTGAATCCGAATGGAAGCATTCAGTGAAGGAAAAATCAATCTGTTGAAGC 600  
DB 541 ATTCATTTGAATCCGAATGGAAGCATTCAGTGAAGGAAAAATCAATCTGTTGAAGC 600  
QY 601 ATCGGTTTATGCGCGGATTAATGATGAAAGATACGCAATCTAAGACAGAAAT 660  
DB 601 ATCGGTTTATGCGCGGATTAATGATGAAAGATACGCAATCTAAGACAGAAAT 660  
QY 661 ACAGATTTTAAAAATTTAGTCAATATTAAGTGAATCAATCTGTTGACCGAGAT 720  
DB 661 ACAGATTTTAAAAATTTAGTCAATATTAAGTGAATCAATCTGTTGACCGAGAT 720  
QY 721 TTTAAAGCTACCAAGCAAAATCTGAGATATTTATCTTTCAGCTCAATATCTCCCT 780  
DB 721 TTTAAAGCTACCAAGCAAAATCTGAGATATTTATCTTTCAGCTCAATATCTCCCT 780  
QY 781 CAAAAAGCTATGGGAAAAAATCACTGTTGGAAGAGATAGAGATATGTAAAGGA 840  
DB 781 CAAAAAGCTATGGGAAAAAATCACTGTTGGAAGAGATAGAGATATGTAAAGGA 840  
QY 841 AATACCAAGCAAAATTAATGATCTGATGTTGGAAGCAAGATGAAATATTAATTT 900

Db 841 AATACCAAGCAATATGTAATCTGATGCTGATTGGAGCAGATGAAATATAAAATT 900  
Qy 901 AGTGGAAAGCTACAAATGGAGATTATTAAAGAAAGGAGAAAGAACTTATAC 960  
Db 901 ACTGGAAAGCTACAAATGGAGATTATTAAAGAAAGGAGAAAGAACTTATAC 960  
Qy 961 ACTCTTTAAGTTTATCAGATGTGGAAGCTTCGTGAAGTAATTAAGAAAGTCATA 1020  
Db 961 ACTCTTTAAGTTTATCAGATGTGGAAGCTTCGTGAAGTAATTAAGAAAGTCATA 1020  
Qy 1021 GGAAGAAATGTGACATTCAGCTGAGAGAAAGAAATTTATATAGCACTTACT 1080  
Db 1021 GGAAGAAATGTGACATTCAGCTGAGAGAAAGAAATTTATATAGCACTTACT 1080  
Qy 1081 AAGCTTGCAAGACACTCTTTTATAGCTTTGTACAGGTTCTATTTCTCTAT 1130  
Db 1081 AAGCTTGCAAGACACTCTTTTATAGCTTTGTACAGGTTCTATTTCTCTAT 1130  
RESULT 8  
AAT78868  
ID AAT78868 standard; cDNA; 5361 BP.  
XX  
AC AAT78868;  
XX  
DT 08-OCT-1997 (first entry)  
XX  
DE P. falciparum liver stage antigen-3 coding sequence.  
XX  
KM Plasmodium falciparum; pre-erythrocyte; liver stage antigen; serum;  
KM prophylaxis; Thai strain; gene organisation; exon; intron; hydrophobic;  
KM glycosyl-phosphatidylinositol membrane anchoring sequence; antibody;  
XX vaccine; immunotherapy; malaria; ds.  
XX  
OS Plasmodium falciparum.  
XX  
FN W09641877-A2.  
XX  
PD 27-DEC-1996.  
XX  
PF 12-JUN-1996; 96WO-FR000894.  
XX  
PR 13-JUN-1995; 95FR-00007007.  
XX  
PA (INSP) INST PASTEUR.  
XX  
PI Drulhe P, Daubersies P;  
XX  
DR WPI; 1997-065464/06.  
DR P-PSDB; AAM24790.  
XX  
PT Plasmodium falciparum poly-peptide(s) and related nucleic acids - derived  
PT from the liver stage antigen-3, useful for malaria vaccine prodn. and  
PT diagnosis.  
XX  
PS Claim 20; Fig 2A-I; 69pp; French.  
XX  
CC This sequence corresponds to the coding sequence for a Plasmodium  
CC falciparum strain KI pre-erythrocytic liver stage antigen-3 (LSA-3)  
CC protein. The gene sequence was isolated by screening a P. falciparum  
CC strain T9/96 library with serum from a missionary treated by prophylaxis  
CC (for strain T6/96 see F93101286). Of 20 clones isolated, clone 7295 was  
CC used to screen a library generated from Thai strain KI. One clone  
CC contained a 6.85 kb insert including the genomic sequence AAT78867. The  
CC gene comprises a 1.8 kb region encoding 3 major blocks of 4 amino acid  
CC repeats and a 3' hydrophobic region corresponding to a glycosyl-  
CC phosphatidylinositol membrane anchoring sequence. The invention relates  
CC to new polypeptides of at least 10 amino acids derived from the LSA-3  
CC polypeptide with the exception of the peptides AAM24791-4. The LSA-3  
CC peptides can be used to raise antibodies and as vaccines for  
CC immunotherapy of malaria  
XX  
SQ Sequence 5361 BP; 2388 A; 431 C; 1169 G; 1373 T; 0 U; 0 Other;

Query Match 1.2%; Score 116.2; DB 2; Length 5361;  
Best Local Similarity 41.7%; Pred. No. 2.5e-12;  
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;  
Qy 5126 GTTCTGATGAAACCGAAGCTTTAGTTAGTGAATTCGAGTTTGAAGAGTAAAGTTCTTCA 5185  
Db 773 GTGTGAAGAAAGATGATAGCTTCAAGTGTGTAAGAAAGTATGATTCAGATATGAAGAAA 832  
Qy 5186 ATGTAGATGCAAAAGATGAAAGAAACCAATAATACAAA--TTGCCGGAATTCGAAATGAG 5242  
Db 833 ATGTAGCTTCAACTGTGTGAAGAAATCGTAGCTCCAAAGTGTGTGAAGAGTGGCTCCAA 892  
Qy 5243 GAAAGCCGCTGAGATTGGAAGCAACAGTTGCTCATCAATATTTGGAAGAAACAATCAGTTA 5302  
Db 893 GTGTGAAGAAAGGTGTAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTGAAGAAA 952  
Qy 5303 TAGCTATTGTAAAAACAGTAATAATTACACCGCGAATGATCAAGATGAAAAATATCA 5362  
Db 953 GTGTAGCTGAAATGTTGAAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAA 1012  
Qy 5363 ATGTGACCTGCAAAAGATTTATCTATGACCAATATCTATGACAGCTCGAAGTTGAGAGCAA 5422  
Db 1013 ATGTGAAGAAATGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAACTGTTGAAGAAA 1072  
Qy 5423 AAGGACCTGTGCAAGAGCTTCTGCAAGTACTACCTTGAATTAAGACAGTTCTCTC 5482  
Db 1073 TTGTAGCTCCAAAGTGTGTAAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAA 1132  
Qy 5483 ATGTGATCAAACTGATATTTGACAAAGATTAGAGGAAGAAATATGAAATTAAGAAA 5542  
Db 1133 ATGTGAAGAAAGGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAA 1192  
Qy 5543 AAGCAAAATGTTAAGTCTAGCTGAAATAAGAGTCAAGTGTCAAAATGCGACAGTGC 5602  
Db 1193 GTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAA 1252  
Qy 5603 TTTCGAGAGCAAGTGTGACAAAGCTGAGAGAGCTGAGAGTGAATTAATTAATTAAC 5662  
Db 1253 ATGTGAAGAAATGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAACTGTTGAAGAAA 1312  
Qy 5663 AAAATACTTGTGACATATTAATAAAATAGTACTCAAAATGTGACGAATGCTTTGTAATA 5722  
Db 1313 TTGTAGCTCCAAAGTGTGTAAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAGAAA 1372  
Qy 5723 GC---AAATCTCATCTATCTATTAACAATTGGAATTGAGCTGAGAGTTGAGAGCTGAG 5779  
Db 1373 ATGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAA 1432  
Qy 5780 GAGCTGAGTACAGAGTTCTGTAGCAGTGAATTAAGATTGTAATTAATTCGATAGCAGAT 5839  
Db 1433 GTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAA 1492  
Qy 5840 TAAATCATGCAAAATATCATGCGAAGGAAATGTGCGAGTTATTACAGAGTGTGATCGCG 5899  
Db 1493 ATGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAA 1552  
Qy 5900 TAAATTCATTAATTAAGCAGGAACAGTGTGAGAGTGTGCCCCGACAGCAATAGAGCTCAA 5959  
Db 1553 TGTAGCTCCAACTGTGTAAGAAATCGTAGCTCCAACTGTGTAAGAAATGTAGCTCCAA 1612  
Qy 5960 CCAAGTGAATGAATTAACAGATCTACAAAGCATATGTAAGAAATTCACAGATTTG 6019  
Db 1613 GTGTGTGAAAGAGTGTGCTCCAAAGTGTGAAAGAAAGTGTGAAGAAATGTGAAGAAA 1672  
Qy 6020 CTAAGAAAGAAACAGATGATTATTAATTAAGAGGCAAGTATGAAGTGTGATGA 6079  
Db 1673 GTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAA 1732  
Qy 6080 AAGATTTCAAAATCTTAATTTAATCAAGAGCTTATCAAGAAAGAAAGAAATTAAGTATA 6139  
Db 1733 ATGTGAAGAAAGTGTAGCT-----GAAATGTTGAAGAAATGTAGCTCCAACTGTTG 1786



```

Db 3936 CAAGATCGAAAAAGTATCTGATTTAAAGATCTTGAGAGATATATTAAAG----- 3988
Qy 8328 TTCTGTAGTATCAGAAAAATGAGAGAAATTCGCCGAGAGAGATTAATCTTCTACAGC 8387
Db 3989 --AAGTAAAGAAATCAAAAGAACTTGAAAGTAAATTTTGAAGATTTAAGAAATTAAA 4046
Qy 8388 ACAAGTAACTAGTAATCCGAGTTCGTTTACGAAAGCAAGATTTAAGAAATATGATTA 8447
Db 4047 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAAGATCATTTTGA AAA 4106
Qy 8448 CACAAAAAAATATTTTCAAGATCAATGCTCTTCTTTAATGATCAAAAGATGAAGC 8507
Db 4107 ATTCGAGAGAGAGCTGAGAAATATAAAGATCTTGAAGCATATATTAAAAAGAGTATC 4166
Qy 8508 GAATATAGATCTTTAGCCGTCGCGTGTGATGCACAGAGAACAAACAAAGCATTTAC 8567
Db 4167 TTCATTAGAGTTGAAGAAAAAAATTTAGAGAGATGACGAAATTTAAAAAGAGAGT 4226
Qy 8568 GAGATCAACAGATTAATCTTACAACTGTAAATGAGAGAAAGTATCTCACTTCGTGC 8627
Db 4227 AGAACATATATTAAGTGTGATGCGCATATAAAGTTTGAAGAGATGATTTAGAGA 4286
Qy 8628 AAAAGCTTTGCTTAAAAATGAAATTTATGGAATGTAAAAAGAACTGAGAGAGCTTACT 8687
Db 4287 AGTAAATGA- TTTAAAAAGAGATATATTAGACATGTTAAAGGAGATATGA- -ATTAGG 4343
Qy 8688 CGAGCCGAGAAACAGACGCCGTTGAAAAATTTATCAAAAGATCTACAGAGCATTTGCTGC 8747
Db 4344 GGAATATGATTAAGAGAAAGTTTAAAGATGTAACAACAAACTTGAGAGAAAGTTGAATC 4403
Qy 8748 AGGAAATTTGGAAATTTGAGATTAATTAGAAACGATTGCAAGATTAATCAATTGTAG 8807
Db 4404 CTTAAAAAGATGTTTATCTAGTGCATTAGGCATGATGAGAAACAAATGAAAAACAAGAA 4463
Qy 8808 AGTCAACGAGAGCGAAGCAAGAGAGCTTTGCGGAAAGATGATTTCTGTGAAAAA 8867
Db 4464 AAAAGCTCAAGACCTTAAGTTGAGAGATATTTAAAAAGAGGTTAAAAAGAGAAC 4523
Qy 8868 TACAATTTTCAGGGGAAACAAATTCATCCATTGAAGATTAAGCCAGAAATTTGTGAACGG 8927
Db 4524 AAAAGAAAAAATTAACAAAAAAGAAAGTAAAGTTGATTTAAGATTAAGAAACAAAAA 4583
Qy 8928 AAGGTAAATAGATGCTTTGAATGAACCTTGATGATCTCAAGAGAAAAAGTGTGG 8987
Db 4584 TGAATATAGTAAAGTTGAAATGAGATGAGATATAGAGAGAGATGTGAAGAGATAT 4643
Qy 8988 CTATGCTGAATTTGATTTGAAATGTTGATTAATATGATTAAGAAAAATGTAGA 9047
Db 4644 AGAAGAGATATAGAGAGAGATTAAGTTGAGATTAAGATTAAGATGAAGATAT 4703
Qy 9048 AGCCAAAAATCGAGAGCATGCTATTGTGAATACTACTGGAAGAAACAGAAATATCAAGCAT 9107
Db 4704 AGGTAGAGCAAGATGAGATTAATTAATAGTCCAAAAAGAGAAACGATTTGAAAA 4763
Qy 9108 TACAAGAGCAAAAGTAAATATTTCTTGAAAGAGAGACGCTGACGCTGCAATATC 9167
Db 4764 GGTAAAGCGAAAAAGAAAAAATTTAGAAAAAAGTTGAAGAGGTGTTAGGCTTAA 4823
Qy 9168 GATATGATCATTTTCCATGAGATGATTTAAAAATTTGGCAAGAGCATATCTTC 9227
Db 4824 AAAACAGTGAAGAGATTAATTAATTTTCAAAAAATTTGATTAAGAGAGTTGATTAAGA 4883
Qy 9228 TCAATTAATTAACCAAAAAATTTCAAAAAATATATTAATTTAGCATCA 9274
Db 4884 AGTATCTTAAGCTTTAGATCAAAAAATGATGTTACTAATGTTTAA 4930

```

RESULT 9  
 ABK5083  
 ID ABK5083 standard; DNA; 5529 BP.  
 AC  
 XX ABK5083;  
 XX

```

DT 07-AUG-2003 (revised)
DT 30-JUL-2002 (first entry)
XX
XX DNA encoding Plasmodium falciparum liver stage antigen-3 (LSA-3) #1.
DE
XX Liver stage antigen-3; LSA-3; vaccine; Th1-inducing adjuvant;
XX malaria parasite; malaria; protein-specific cytotoxic T cell response;
XX gamma-interferon; gene; ds.
OS
XX Plasmodium falciparum.
FH
FH Key
FH CDS
FT 1..5529
FT /tag= a
FT /product= "LSA-3"
FT /note= "Liver stage antigen-3"
FT /transl_except= (pos:5524..5526, aa:lys-glx)
FT 1..198
FT /tag= b
FT /number= 1
FT 199..366
FT /tag= c
FT /number= 1
FT 367..5529
FT /tag= d
FT /number= 2
PN
XX EP1201250-A1.
XX
XX 02-MAY-2002.
XX
XX 25-OCT-2000; 2000EP-00203724.
XX
XX 25-OCT-2000; 2000EP-00203724.
XX
XX 25-OCT-2000; 2000EP-00203724.
XX
XX (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX (INSP ) INST PASTEUR.
XX
XX Cohen J, Drullhe P;
XX
XX WPI; 2002-373883/41.
XX
XX P-PSDB; AAU96699.
XX
XX Vaccine for treatment or prevention of malaria, comprises liver stage
XX antigen and adjuvant that induces Th1 response.
XX
XX Disclosure; Page 16; 56pp; English.
XX
XX The invention describes a vaccine comprising a Th1-inducing adjuvant (I)
XX and a protective liver-stage antigen (LSA), or its immunogenic fragment,
XX from a human malaria parasite, where LSA is a fragment of LSA-3, then (I)
XX is not montanide. The vaccine is useful for treating or preventing
XX malaria, specifically where caused by Plasmodium falciparum. The vaccine
XX is formulated in SBA52 adjuvant, an oil-in-water emulsion containing QS21
XX and 3D-MPL (de-O-acylated monophosphoryl-lipid A, the combination of
XX which results in strong induction of a circumsporozoite protein-specific
XX cytotoxic T cell response, which is not generally induced by vaccines
XX based on recombinant proteins, and synergistically increases production
XX of gamma-interferon. This sequence encodes the liver-stage antigen-3 (LAS
XX -3) from the Plasmodium falciparum strain KI used to develop the
XX recombinant proteins for the vaccine. (Updated on 07-AUG-2003 to correct
XX OS field.)
SQ
XX Sequence 5529 BP; 2473 A; 435 C; 1174 G; 1447 T; 0 U; 0 Other;
XX
XX
XX Query Match 1.2%; Score 116.2; DB 6; Length 5529;
XX Best Local Similarity 41.7%; Pred. No. 2.5e-12;
XX Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;
Qy 5126 GTTCTGATGAACGAGCTTTAGTATGATTTCTGAGATTGAAGAGTAACTTCTTCA 5185
Db 941 GTGTTGAAGAAAGTATAGCTTCAAGTGTGATGAAGATTAATTAATTAAGAGAA 1000

```

Qy	5186	ATGTAGATGCAAAAGATCAAAAAACAATAATACAA---TTGCCGAAATATGCAATATGAG	5242
Db	1001	ATGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGTATGAAGATGTGCTCCA	1060
Qy	5243	GAAAAGCGGCTGAGGTTGGAGCAACAGTGTCTCTAACAAATATGTGAAAACAATCAAGTTA	5302
Db	1061	GTTGTTGAAGAAAGTGTAGAGAAAATGTTGAAGAAAGTATAGCTGAAAATGTTGAAGAAA	1120
Qy	5303	TAGCTATTGTAAAAAACGTAAAAATTACAACGGCAATGATCAAGATAGAAAAAATATCA	5362
Db	1121	GTTATAGCTGAABAATGTTGAAGAAAGTATAGCTGAAAATGTTGAAGAAAGTATAGCTGAAA	1180
Qy	5363	ATGTGACTGCAAAAGATTATATCTATGACCAATCTATAGCAGTCCGAGTTGGAGAGCA	5422
Db	1181	ATGTTGAAGAAATCGTAGCTCCACTGTGGAAGAAATCGTAGCTCCAACGTGTTGAAGAAA	1240
Qy	5423	AAGGAGCCTCTGTCCAGAGACCTTCTCGAAGTACTACCTGAAATATACAGTTTCTTCTC	5482
Db	1241	TTGTATGCTCCAAAGTGTGTATGAAGATGTGGCTCCAAAGTTTGAAAGAAAGTATAGAGAAA	1300
Qy	5483	ATGTTGATCAAACTGATATTTAGCAAAAGATTATAGAGAAAGAAATATAGAAATATAGAAA	5542
Db	1301	ATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTATAGCTGAAAATGTTGAAGAAA	1360
Qy	5543	AGGCAAAATGTTATGTTCTAGCTGAAAAATACGAGTCAAGTGTGCACAAAATGCGACAGTGC	5602
Db	1361	GTTATAGCTGAAGAAATGTTTAAAGAAAGTATAGCTGAAAATGTTGAAGAAAGTATAGCTGAAA	1420
Qy	5603	TTTCCGAGCAAGTGTGACAGCTGACATGAGAGCTGGAGTGTGCAATTATATTAATTTATAC	5662
Db	1421	ATGTTGAAGAAATCGTAGCTCCAACTGTGGAAGAAATGTATGCTCCAACGTGTTGAAGAAA	1480
Qy	5663	AAATATCTTGTGCACATATTAATAAATATAGTACTCAAAATGTACAAATATGCTTGTGTAATA	5722
Db	1481	TTGTATGCTCCAAAGTGTGTATAGAAATGTGGCTCCAAAGTGTGAAGAAAGTGTATAGAAA	1540
Qy	5723	GC---AAATCTCATTCATCTATTAATAACAATTTGAATTTGAGCTGAGTTGAGACTGAG	5779
Db	1541	ATGTTGAAGAAAGTGTATGTAATAATGTTGAAGAAAGTGTATGTAATAATGTTGAAGAAA	1600
Qy	5780	GAGCTGTAGTGTACAGGTTCTGTAGCAGTGAAATAGATTGTATATATATACATATACAGAT	5839
Db	1601	GTTATGCTGAAAATGTTGAAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTATGTAATA	1660
Qy	5840	TAAATCATGCAAAAATCACTCGGAAGGGAATATGTCGAGTATTAACAGATCTGATCGG	5899
Db	1661	ATGTTGAAGAAAGTGTATGTAATAATGTTGAAGAAAGTGTATGTAATAATGTTGAAGAAA	1720
Qy	5900	TAAATGCTAATTTATGCAAGAAACAGTGTCTGAGTGTGAGTGTGCCCCGTGCAATATAGACCTCAA	5959
Db	1721	TGTTATGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATTTGTACTCCA	1780
Qy	5960	CCAGTGTGATGAAAATTAACAGGATCTACAAAGCATATGTATTAAGATTCTACAGTATTG	6019
Db	1781	GTTGTTGTGAAAAGTGTGCTGCCAGTGTGGAAGAAAGTGTGAAGAAATGTTGAAGAAA	1840
Qy	6020	CTAAGAGAAACAAATGATTTATATTAATCTACTCAAGGCAAGTGTATTAATGTGTAGATA	6079
Db	1841	GTTATGCTGAAAATTTGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTATGCTGAAA	1900
Qy	6080	AAGTATTTCAAAAATCTTATATATTAACGAAGCTTATCACAAAAAAGAAAAATTAAGTATA	6139
Db	1901	ATGTTGAAGAAAGTGTAGCT-----GAAAATGTTGAAGAAATGTGTACTCCACTGTGTG	1954
Qy	6140	AAAAAGATTTGTTATCCAAATAGTTCACTCATATCTTAAATCTTTATTTGCGAATG	6199
Db	1955	AAGAATATGTAAGTCTCAACTGTTGAAGAAATGTATGCTCCAAAGTGTGTA--GAAAGTGTG	2011
Qy	6200	CGGCTGTTCACAGAACGCGAGTGTGCAAGAACTGTTAATATCAAAAGGTTATATGAG	6259
Db	2012	TGGCTTCCAAGTGTGTAGAAAATGTATGAAGAAATGTTGAGAAAGTGTATGCTGAAAATG	2071
Qy	6260	AAACGAAAGCTCTGTAGAAAATTTCTATATTAATGCAAAACATTATTTCTGTAAATATCAG	6319

Db	2072	TTGAGAAAGTGTAGCTGAAAAATGTTGAAGAAATGTAGCTGAAAAATGTTGAAGAAATCG	2131
Qy	6320	GAGATTACAGAAATTCATCGAGTAGTAGGTTCTGTGGTGTGGTGAATGTAGAG	6379
Db	2132	TAGCTCAACGTGTGAAGAAATGTAGTCTCAACTGTGTGAAGAAATGTAGCTCAAGTG	2191
Qy	6380	TAGAGCTTCTTCTGATACCAATTTATTAAGG--AAATCCAAACAAGTTGGA	6436
Db	2192	TTGTAGAAAGTGTGGCTCCAAAGTGTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTG	2251
Qy	6437	AAACTACAAATGTGATGAAGGTTTGGAGAAAGCTGAAATTAACGAGATTTCTAAGC	6496
Db	2252	TAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGTGAATATG	2311
Qy	6497	AAGCAATTTCCCTCTTTTGGAGTCGAGTCGACAGCCGGGTAGGAGCCGAGTGCAG	6556
Db	2312	TTGAGAAAGATGTAGCTGAAAAATGTTGAAGAAATGTAGTCTCAACTGTTGAAGAAATCG	2371
Qy	6557	GAACCGTTTCCGTAATCAATTTGCAGAGAAAGCGAAGTAGATGTGAAGAACCAAGA	6616
Db	2372	TAGCTCAACGTGTGAAGAAATGTAGTCTCAAGTGTGTAGAAAGTGTGGCTTCCAAAGTG	2431
Qy	6617	TTTTGTGTAAAAAAGCTGAGATTACACAAACGTTATGTTCTGTGCAATTTGGAATG	6676
Db	2432	TTGAGAAAGTGTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTG	2491
Qy	6677	CCGCAGTCCGAGTGGCTGCAAAAAGAGCTGGAAATTTGAGCAGCAGTGGCAGTTTCCAAAG	6736
Db	2492	TAGCTGAAAAATGTTGAAGAAAGTGTAGCTGAAAAATGTTGAAGAAAGTGTAGCTCCAACTG	2551
Qy	6737	ATGATCAAAACACAGAGCAAGTAGTGAAGAAATTTCTAAATTT--ATGACTGAAACAAGT	6793
Db	2552	TTGAAGAAATGTAGTCTCAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGTG	2611
Qy	6794	TAGATGTATATGCAAGAAATGAGTAAATCAGTACTCGAATCGGTTCCGCGGAGCTG	6853
Db	2612	TTGCTGAAAAAGTGTGCAACAAATTTATCAGCAACTTTTAAAGTAATTTATTAAGTGTGA	2671
Qy	6854	GAATCTTTCAGCCGAGTAGTCTGAGTGTGTTCTGTCAATATATTTGCCAATTAAGTAG	6913
Db	2672	TCGAACTGAGGAATTAAGACAGCTATTTAAATGAGATGAGAAAGTAAAGAAATG	2731
Qy	6914	AAACGATATGCATATAGTACTTTACACTTCTTACTGATGATTAATGTAAGACTCTTA	6973
Db	2732	TAGTCACACAATCTGAAAAACGTAGAAAGAACTACAGCTGAAGATGTACTACTTTTA	2791
Qy	6974	ATAAAAATTTGCAATTCCTTGACAGCCGGTGAAGAGCCGCGGTCTTGACAGACTTACCG	7033
Db	2792	GTAACATATTAGAGAGATACMAAAAAATCTATTACTATGATATCTATAGAGAAAAAT	2851
Qy	7034	GAGGTGTTCTGTTAACACTAATAATGTTCTGTGATAGTCTCGAGTTCACATTAATCTG	7093
Db	2852	TGAGAAAGCTCCAGAAATATTTAAAGTGTGCGCTTTGAAAAATACCCAAAGTGAAGG	2911
Qy	7094	ATTTGACTTCCGTAACGAGA--AAAAGTAATGTAAACGCCAAAGAGGAAAAATATTA	7156
Db	2912	AAAAGAAAGATTAATAGATGTAATTTGAAGATTAAGAAAGAGAGTGCTACACTTTAA	2971
Qy	7151	AGCAACAGCAGCAATTCAGGAATTCGAGAGCAGCAATCGAGCCAAATGCTTTGTGA	7210
Db	2972	TAGAAACTGTGAAACAGGCAGAGAAAGACGCGMAATTAATTAACGAAATATTTGAAA	3031
Qy	7211	ATTAATTTTGAACAGCTGTAGAGATGAAAAAATTTGGAAGAAAAAGAAACAGAGTTT	7270
Db	3032	ATTTAGAAAGAAATGTAGTGAAGATTAAGAAATGTTGCGAGAGATTTAGAGAAATTA	3091
Qy	7271	TAAAAAATTTAGACGAAGTTAAACAAGAACAAATTAAGTAATGATCTACGAAA	7330
Db	3092	ACGAAACTGTATTTAATCTGATTAATAGTAAAGTAGAGAAACAGTAGAAATTTAGCGAG	3151
Qy	7331	AAATTTTCAATCAGCAGTATTTCTACAGAAATACCTTCTGTAAAGCGATAGAGAG	7391





QY 6437 AAACCTACAAATGCTGTATGAAGTTTCGAGAAAGAAAGCTGAAATTTACAGACAGATCTTAAGC 6496  
DB 2338 TAGCTGAAAATGTTGAAAGAAAGTGTAGCTGAAAATGTTGAAAGAAAGTGTAGCTGAAAATG 2387  
QY 6497 AAGGAATTTCTCTTTTGGAGTCGAGTCGACAGCCGCGGATAGAGCCGAGAGTCGAG 6556  
DB 2388 TTGAAGAAAGTGTAGCTGAAAATGTTGAAAGAAATGCTAGCTCCAACTGTGGAAGAAATCG 2447  
QY 6557 GAACCGTTTCGCTAAATCAATTTGACAGAAAGACGAAAGTAGATGTGAAAGAAAGAGA 6616  
DB 2448 TAGCTCCAACTGTGAGAAATTTGTAGCTCCAAAGTGTGTAGAAAGTGTGCTCCAAAGT 2507  
QY 6617 TTTTGGTAAAAAGCTGAGATTTACAGCAAAACGTTATAGTCTGTGTCATTTGAAATG 6676  
DB 2508 TTGAAGAAAGTGTAGAAAGAAATGTTGAAAGAAAGTGTAGCTGAAAATGTTGAAAGAAAGT 2567  
QY 6677 CGGAGTCGAGTGGCTGCAAAAAGAGCTGGAATTTGAGACACAGTCGAGATTTACCAAG 6736  
DB 2568 TAGCTGAAAATGTTGAAAGAAAGTGTAGCTGAAAATGTTGAAAGAAAGTGTAGCTCCAACTG 2627  
QY 6737 ATGAATCAAAACAGAGACGAAAGTAGAAAATTTCTAAAATTT--ATGACTCGAAACAAAGT 6793  
DB 2628 TTGAAGAAATTTGAGCTCCAAAGTGTGAAAGAAAGTGTAGCTCCAAAGTGTGAAAGAAAGT 2687  
QY 6794 TAGATGTAAATAGCAAGAAATGAGATTAATCAAGTACTGGAATCGGTTCAAGCCGAGCTG 6853  
DB 2688 TTGCTGAAAACGTTGCAACAAATTTATCAGCAATCTTTTAAGTAATTTATAGGTGGTA 2747  
QY 6854 GAATCTTGACAGCCGAGATCTGGAGTGGTTCTGTGCAATTAATTTGGAATTAAGTAG 6913  
DB 2748 TCGAAACTGAGAAATTAAGACAGTATATTAATAGATAGAGAGATTAAGAAAGAAATG 2807  
QY 6914 AAACAGATATCATCTACTACTTTTACACTCTTCTACTGATGTAATGTAAGTAAAGCTCTTA 6973  
DB 2808 TAGTCACCAATATCTAGAAAACGTAGAGAAAGTAACTACAGCTGAAAGTGTATCTACTTTTA 2867  
QY 6974 ATTAATTTTCGAATTCCTTGACAGCCGCTGAGAGAGCCGAGCTCTTGACAGATTTACCG 7033  
DB 2868 GTTACATATTTAGAGAGATACAAAGAAATATCTATGATATCTATAGAGGAAAT 2927  
QY 7034 GAGTGGTTCTGTAACTATTAATAGTCTGTGATAGCTGAGTGCATTAATCTG 7093  
DB 2928 TTAGAAAGACTCCAGCAAAATGTATTAAGTGCCTTTAGAAAATGCCAAATGGAAGG 2987  
QY 7094 ATTTGACTTCGCTACGAGA---AAAAGTAATGTAAACGCAAAAGAGAAAATATATTA 7150  
DB 2988 AAAAGAAAGATTAATAGATGTATTTGAGAGAGTAAAGAGAGTGCCTACACTTTAA 3047  
QY 7151 AGCAAAACAGACCAAAATGACAGATTCGAGAGACAGCAATCGAGCCAAATGTTCTGTAA 7210  
DB 3048 TAGAAAATGCTGAAACGCGCAAGAAAGAGCGCAATCAATTAACGGAATAATTTGAAA 3107  
QY 7211 ATTAATTTTGGACAGCTGTAGAAAGATAGAAAATTTCTGAAGAAAAGAAACAGAGTTT 7270  
DB 3108 ATTTAGAGAAATGACAGTAAAGTATGAAAATGTTGACAGAAATTTTGAAGAAATTA 3167  
QY 7271 TAAAAACTTTAGACGAAAGTTAAACAAAGAACAGATTAAGAAAGTAATGATGCTACGAAA 7330  
DB 3168 ACGAAACGTATTTATATCTGTATTAAGATAAGTAAAGTAAAGAGAAACAGTAAATTTACCGAG 3227  
QY 7331 AAATCTTACATCAGAGATTTCTTACAGAAAGATCTTCTGTAAAGCGGATAGAGAG 7390  
DB 3228 AAAAGTTTGAAGAAACATGAAATGGAATTAAGCAATTTTGTAGAAAATATTTGTAATGTAA 3287  
QY 7391 ATACTCAGGAGAAAGAAATTAAGCATTTGTGAAGCTTCTGATATTAATTTGAAAAATG 7450  
DB 3288 AAGGAATACAGAAAATTTATTAACAGGTATG-----TTTGAAGTATAGAAACACAGTA 3341  
QY 7451 TAGATATTTCAACAGAGACAAGAAATATATCATCTCTAGTGTGTTGGAACTGCG 7510  
DB 3342 TAGATATTCATCAAGAGAAAGGTTGATTTTGAATGAAAAATGTGTAGTTCGATTTTNG 3401  
QY 7511 GTCTTGCTTCCGATCAGGAACAGTGGCAGTTACAAATATTTAAAGAAATTCGCGAGTTA 7570

DB 3402 ATTAATATGAAAAATATGAAAAAGGTTTATTAATTAATTTAGAAAATTTTCAAGTACTG 3461  
QY 7571 CTGTTGAAAATTTCTTTGTGAAAGAGCTGAAAAAGTAAATGTTAGATCGGAT----- 7625  
DB 3462 AAGGTGTCAAGAAACGTAACTGAAATGTAGAACAAATATATATGTGTGATGTTGATG 3521  
QY 7626 ----TACAGAAATGTTGCTTTAACAGCATATCAAGGTCTGTAGAGACATTTGGAAATG 7681  
DB 3522 TTCTGCTATGAAAGATCAATTTTATAGAAATTAATTAATGAGGACGAGGTTGAAAGAA 3581  
QY 7682 GAGCTGCTATGCAAGATTAATTTCTAATGGAAGATCAATATCTGATTTAATAATCTTA 7741  
DB 3582 TGTTTTAAATTTGGAAAGATGATTTTAAAGTGAAGTGAATTAATCTGTAGAGAA 3641  
QY 7742 AGCTATTAGAAAAAATTTGATGTTATTTGTAAGATTAATTCGAATTTGAGAGCGAG 7801  
DB 3642 TTAAGATGAACCGGTTCAAAAAGAGTAGAAAGAAACGTTAGTATTTATGAAAGAA 3701  
QY 7802 CAAGAAGATTAACGTAAGACCGTAGCTGCGGAGCCATTAATCTCAAAAGCAAGAAATG 7861  
DB 3702 TGAAGAAATATATGTATATGATTAAGAGAAAGAAAGAAAGATTTAACAGCAAGATGA 3761  
QY 7862 AAATGAATTCAGAGTTGAAATTTGAGAAAGATTTTCAATGAAGAAATGAGTAACTA 7921  
DB 3762 TAGATGAGTAAAGATCCATAGAAATATCTTCAGATTTCAAGAGAAAGAACTGAATCTA 3821  
QY 7922 GCCCTTTAAAGGAATTTGAAAGAAATCAATGTCAAAAGTGGAAAAAGAAACAGAGTA 7981  
DB 3822 TTAAGATTAAGAAAGAAAGATGTTTCACTAGTGTGAAAGAAAGTTCAACACATGATATG 3881  
QY 7982 CTGCTGAATCTCAAGAGACT-----TCTGTAGAGACAGTAGAGGCGCAG 8027  
DB 3882 ATGAAAGTGTGAGAAAGTTTAGATTAAGAAATATGAAAGAGAGATTAATGAAAGATG 3941  
QY 8028 AATTAATTTCCAGAGCAAAAGATGCCGAAAGCTTTATTTGAAAGTAAATTAATCCGG 8087  
DB 3942 CTGTTGAAATTAAGACATTTCTAGCAAACTTATTTGAAGAACTCAAGAGTTAAATGAG 4001  
QY 8088 AAGAAGTATTTTTCATGAGATTAATGTGAATATGGAAGCAACATTAATGAAGTAAG 8147  
DB 4002 TTAGAGCGATTTAATTAAGAAATGTGAAGAAATTAAGAAATTTAGAAAAAGC--ATTAATC 4059  
QY 8148 AGCAGTTTCTAAGACAGTACAGGTTCTGTATTTGGAGAGATTTGAGTCAACCAAGCAGA 8207  
DB 4060 AAGAGATTTCTAAGAAATTAATAGATGCAAAAGATGATCATTTAGAAAAGTATTTGAAGA 4119  
QY 8208 AGCTATGCTCGAGGTAAACATATGTAGAAAGTTGAGAGAAATTTGTTCAAGACAA 8267  
DB 4120 GGAACATGATATTAACGACGAGGTTGAGTGAATTTGTAGAAATTAAGAAAGTGTGAAAGAG 4179  
QY 8268 TCGATTAATGCAATTTCTAAGTGAAGGTTTGAATGAATTAAGTAAGTAAGTAAATC 8327  
DB 4180 CAGATGCAAAAAGTATGTTGTTTAAAGATCTTGAAAGAAATATTAATTAAG----- 4232  
QY 8328 TTCTGTAGTATCAGAAATGAGAGAGATTTGCCGAGACAGAGTGAATCTTACAGC 8387  
DB 4233 --AAGTAAGAAATCAAAAGAACTGAAAGTGAATTTTAAAGATTTAAGAAATTA 4290  
QY 8388 ACAAAGTATATCTGAATCCGATGCTTTACGAAAGCAAGATTTGAAAATTAATGATTA 8447  
DB 4291 AACTATTGAAACAGATATTTTGAAGAGAAAAAGAAATAGAAAAAGATTCATTTTGAAGA 4350  
QY 8448 CACAAAAAATATATTTTCAAGAGCAATGCTCTGTTAAATGATACAAAGAAATGAAGC 8507  
DB 4351 ATTCGAAGAAAGAGCTGAAGAAATTAAGATCTTGAAAGCAATATATTTAAAGAAATATC 4410  
QY 8508 GAATATGAATCTTTAGCGGTAGCGGTGTCATGCAAGAAACAAACAAAGCATTTAC 8567  
DB 4411 TTCAATTAAGATTTGAAGAGAAAAAATTTGAAGAGTACAGAAATTAAGAAAGAGGT 4470  
QY 8568 GAGATCAAAACAGTTAATCTTACACTGTAAATGAGAGAAACGTATCTCAATTCGTGC 8627

Db 4471 AGAATATATTAAGTGGATGCGCATATAAAGGTTTGGAGAAGATGATTGAAGA 4530  
Qy 8628 AAAAGCTTGGCTAAAAATGAAAATTATGAAATGTAAAGAACTGAGAGCTTGT 8687  
Db 4531 AGTAGATGA-TTTAAAGAGAGTATATTAGACATGTTAAAGGAGATATGA--ATTAG 4587  
Qy 8688 CGGACCGGAAACAGAGCCGTGAAAATTATCAAAAGAGTACTACAGAGCTTGGTGC 8747  
Db 4588 GATATGATTAAGAAAGTTAGAGATTAACAACAACTTGGAGAAAGCTTGAATC 4647  
Qy 8748 AGGAATTGGAAATTGGAGATAAATTAGAAAACGATTCAGAGATTAATTCGATTGTAG 8807  
Db 4648 CTTAAAGAGTGTATTATCTAGTCATTAAGCATGATGAGAAACAAATGAAACAAAGAA 4707  
Qy 8808 AGTCAACGAGACGAAACCAAGAGCTTGTGCGAAAGAAATGTTCTGTGAAA 8867  
Db 4708 AAAAGCTCAAGACCTTAAGTTGAGAAAGTATTAATAAAGAGAGTTAAAGAAACC 4767  
Qy 8868 TACAAATTTGAGGGGAAACAAATTCATCCATTGAAGTAAAGCAGAAATTTGGAAACGG 8927  
Db 4768 AAGAGAAAAATTAACAAAAAGAAAGTAAAGTTGATTAAGATTAAGAAACAAAGAA 4827  
Qy 8928 AAGTGAATGTAGATGCTTTGAAATGAATCTTGATGATGATCAAGAAAGAGTGG 8987  
Db 4828 TGAATATGAGAAAGTTGAATGAAATGAAAGTATTAAGAAAGATGTAGAAAGATAT 4887  
Qy 8988 CTATGCTGAATTTGATTTGAAATGTTGATTAATTAATGATTAGAAAAATGTAGA 9047  
Db 4888 ABAAGAGATTAAGAAAGATTAAGATTAAGTGAATATGATGATGATGATGATGAT 4947  
Qy 9048 AGCCAAATCGGAGACATGCTATTGTAGAACTACTGAGAAACAGAAATTCAGCACTT 9107  
Db 4948 AGGTGAAGCAAGATGAAGTATTAAGTTCAAAGAAAGAAACGCAATGAAAA 5007  
Qy 9108 TCAAGAGCAAAAGTAAATATCTTGTGAAAGAGACGCTGACGTGACATGATATC 9167  
Db 5008 GGTAAAGCGAAAGAAAGAAATTAAGAAAAAGTTGAAGAGGTGTTAGTGTTTAA 5067  
Qy 9168 GAATGACATTTCCATGAGATGATTAATAAAATTTGGCAAGCAGATGATCTTC 9227  
Db 5068 AAAACGCTAGACGAGTATGAAATATGTTCAAAAAATGATTAAGAAAGTATGAAGA 5127  
Qy 9228 TCAATTAATACCAAAATTCAAAAATTAATTAATTAATTAATTAATTAATTAAT 9274  
Db 5128 AGTATCTAAAGCTTTAGAAATCAAAAAATGATGTTACTAATGTTTAA 5174

RESULT 11  
AAT05868 standard; DNA; 3399 BP.  
ID AAT05868;  
XX  
AC AAT05868;  
XX  
DT 27-AUG-2003 (revised)  
DT 14-AUG-1996 (first entry)  
XX  
DE Chicken leucocytozoan DNA encoding immunogenic protein for vaccines.  
XX  
KW Chicken leucocytozoan; immunogen; recombinant vaccine; protection;  
KW immunisation; vaccination; ss.  
XX  
OS Leucocytozoan sp.  
XX  
XX Key Location/Qualifiers  
PH 1..3399  
FT CDS  
FT /\*tag= a  
FT 1150..33218  
FT misc\_feature  
FT /note= b  
FT insert in a recombinant vaccine against chicken  
FT leucocytozoan disease"  
XX  
XX JP07284392-A.  
PN

XX 31-OCT-1995.  
PD 19-APR-1994; 94JP-00080643.  
XX 19-APR-1994; 94JP-00080643.  
XX 19-APR-1994; 94JP-00080643.  
XX (DOBU-) DOBUTSUO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.  
PA (KITA) KITASATO KENKUSHO SH.  
XX  
XX WPI: 1996-006311/01.  
DR P-PSDB; AAR97866.  
XX  
XX Chicken leucocytozoan immunogenic protein - used in a recombinant vaccine  
PT against chicken leucocytozoan disease.  
XX  
XX Claim 6; Page 6-9; 35pp; Japanese.  
XX  
XX AAT05868 encodes a chicken leucocytozoan immunogenic protein, this DNA or  
CC a fragment of it can be used in a recombinant vaccine to immunise against  
CC chicken leucocytozoan disease. The DNA is used in a vector and  
CC operatively linked to an expression regulatory sequence as in standard  
CC practice. (Updated on 27-AUG-2003 to correct OS field.)  
XX  
XX Sequence 3399 BP; 1577 A; 508 C; 798 G; 516 T; 0 U; 0 Other;  
SQ

Query Match 1.1%; Score 104.2; DB 2; Length 3399;  
Best Local Similarity 43.3%; Pred. No. 4.8e-10;  
Matches 748; Conservative 0; Mismatches 963; Indels 16; Gaps 5;

Qy 7110 AGAAAGTAAATGATGACGCAAAAGAGAAAAAATTTAAAGCAAGCAGCAAAATGC 7169  
Db 1461 AAACCAACATTTGAAACAGAAATTAACAATCTATGAAACGATTAATAAACAATGA 1520  
Qy 7170 AGAAATCGAGAGAGCAGCAATCGAACATCTTTGTAATTAATTTTGGACAGCTGT 7229  
Db 1521 TGAAGAACGAAACACACATTAATGAAGAAATTAATTAATGATGATGATTTACCA 1580  
Qy 7230 AGAATGAGAAATTTCTGAAAGAAAGAAACAGAAAGTTTAAATCTTGAAGCAAGT 7289  
Db 1581 ACTTAATGTAGAAAGATCAGATGAACAGAAATTAACATATGAAATGATTAATGACAT 1640  
Qy 7290 TAAACAAGAAACAAGTAAATAAGTAAATGATCTACGAAAAAATCTTCAATCAGCAG 7349  
Db 1641 ACAGAGAGCATTGAAAGAAATGTAACATGAAAGAAAGAAAGAAAGTAAACATGAAGA 1700  
Qy 7350 TATTTCTACAGAAATCTTCTGTAAAGCGATAGAGAGATCTCAGAGAGAGAAAT 7409  
Db 1701 AATAGAAAAAGAAAGCAGTAAAGAAATTAATGAAAGAAAGAAAGTAAACACA 1760  
Qy 7410 TAAAGCATTTGAGACATTTCTGATTTATTGGAATAATGTAATTTTCAACAGAGA 7469  
Db 1761 TGAAGAAATTAAGAAAGAAAGACATGAAAGTAAATTAATCAATGAAAGAAAGAAAGT 1820  
Qy 7470 CAAAGTAAATATCACTTCTACTGCTGTTGGAACTGCAAGCTTGTCCGATCAGG 7529  
Db 1821 AACACATGAAAGAAAGAAAGAAAGAGCATGAAGAAATTAATCAATGAAAGAAAGAAAG 1880  
Qy 7530 AACAGTGCAGATT--ACAATATTTAAAGAAATTCGAGATTAATCTGTTAAATTTCTT 7587  
Db 1881 AGAAGTAAACATGAAAGAAATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1940  
Qy 7588 GTGAAAGCAGCTGAAAGAAAGTAAATGTTAGATCGATTTTACAGAAATGTGCTTTAA 7647  
Db 1941 AAAAGAAAGAAAGTAAACATGAAAGAAATTAAGAAAGAAAGAAAGAAAGAAAGTAA 2000  
Qy 7648 GCATATCAAGGTCTCTAGAGATGGAATTAAGAGTGCCTATCAGAAATTAATTTCT 7707  
Db 2001 GGAAGAAAGAAAGAAAGTAAACATGAAAGAAATTAAGAAAGAAAGAAAGAAAGAAAG 2060  
Qy 7708 AATGAAAGTCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7767  
Db 2061 ACATGAAGAAAGAAAGAAAGTAAACATGAAAG--AAATGAAAGAAAGAAAGCAATGA 2118

```

QY 7768 ATTGTAAGATTAATCGAATTGAGCGGAGCAAAAGATTAAACCTAGACCGGTA 7827
DB 2119 GAAGTAATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2178
QY 7828 GCTGCCGAGGCAATTAATCTCAAAAGCAAGAAATGAAATTAATTCAGAGTTGAATTGAG 7887
DB 2179 CATGAGAGAGTAATACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2238
QY 7888 AAGAGTAATTTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7947
DB 2239 AAGAGTAATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2298
QY 7948 ATCAATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8007
DB 2299 ATACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2358
QY 8008 GGAGCGAGTACAGAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8063
DB 2359 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2418
QY 8064 -TTTGAGAGTATGATCAAAATCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 8121
DB 2419 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2478
QY 8122 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8178
DB 2479 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2538
QY 8179 ---TTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8235
DB 2539 CATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2598
QY 8236 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8295
DB 2599 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2658
QY 8296 GGTGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8355
DB 2659 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2718
QY 8356 ATTGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8415
DB 2719 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2778
QY 8416 TTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8475
DB 2779 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2838
QY 8476 GCTCTTGCTTTAATGATACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8535
DB 2839 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2898
QY 8536 GTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8595
DB 2899 GTTACACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2958
QY 8596 GTTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8655
DB 2959 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3018
QY 8656 GGAAGATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8715
DB 3019 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3078
QY 8716 TATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8775
DB 3079 TATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3138
QY 8776 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8832
DB 3139 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3185

```

```

RESULT 12
AD272262
ID AD272262 standard; DNA; 1701 BP.
XX
AC AD272262;
XX
DT 14-JUL-2005 (first entry)
XX
DE Plasmidium falciparum MSP-3-like gene, MSP-3-6 SEQ ID NO: 11.
XX
KM Nucleic acid vaccine; plasmidium falciparum infection; antimalarial;
XX infection; merozoite surface protein 3-like gene; MSP-3-6; chromosome 10,
XX gene; db.
XX
OS Plasmidium falciparum.
XX
FH Key Location/Qualifiers
FT CDS 1..1701
FT /tag=a
FT /product="Plasmidium falciparum MSP-3-6 protein"
XX
PN EP1526178-A1.
XX
PD 27-APR-2005.
XX
PF 24-OCT-2003; 2003EP-00292673.
XX
PR 24-OCT-2003; 2003EP-00292673.
XX
PA (INSP) INST PASTEUR.
XX
PI Dru11he P;
XX
DR MPI: 2005-323987/34.
DR P-F8DB; AD272263.
XX
PT Novel MSP-3-like family genes located on chromosome 10 of Plasmidium
PT falciparum, which encode proteins useful for preparing vaccine
PT compositions against malaria.
XX
PS Claim 6; SEQ ID NO 11; 137bp; English.
XX
CC The present invention relates to the protection against malaria. More
CC particularly, the invention pertains to a family of MSP-3 (merozoite
CC surface protein 3)-like genes (MSP-3-1, MSP-3-2, MSP-3-3, MSP-3-4, MSP-3-
CC 5, MSP-3-6, MSP-3-7 and MSP-3-8) located on chromosome 10 of Plasmidium
CC falciparum, highly conserved in P. falciparum strains, simultaneously
CC expressed in P. falciparum at the erythrocytic stages and encoding
CC proteins which have a Asn-Leu-Arg-Asn or Asn-Leu-Arg-Lys signature at
CC their N-terminal extremity and which are located at the merozoite
CC surface. The characterization of this gene family enables the definition
CC of immunogenic and vaccine compositions against P. falciparum. The
CC present sequence is the P. falciparum MSP-3-6 gene.
XX
SQ Sequence 1701 BP; 804 A; 94 C; 378 G; 425 T; 0 U; 0 Other;
QY Query Match 1.0%; Score 98.6; DB 14; Length 1701;
Best Local Similarity 42.8%; Pred. NO. 4.8e-09;
Matches 610; Conservative 0; Mismatches 809; Indels 6; Gaps 2;
DB 7052 CTATTAATAGTTCGTGATAGCTGAGTTCAGATTAATCTGATTTGACTTCGATACGAG 7111
DB 110 CTTTAAAGAGCAATTTTAAAGAGATGATTTTAAATTAATCTTATTTTAAAGAGAGAGCA 169
QY 7112 AAAAAGTAATGTAAACGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7171
DB 170 ATAATTTTAAATATGAGAAATGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 229
QY 7172 GAAATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7231
DB 230 CTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 289

```

```

Qy 7232 AAGTAGAAAAATTCGAGGAAAAGGACAGAGTTTAAAACTTTAGACGAGTTA 7291
Db 290 TCTTACTGAAAAATPAAGAACAAAAATGAGAGATCCATGAAAAATGAGATGTTAA 349
Qy 7292 ACAAGAAACAAGATPAAAAAGTAAATGATGCTACGAAAAAAATCTTCAATCAGCAGSTA 7351
Db 350 ATGATGAGAGAAAGTAAAAACAGAAATGTATCTGAAAAAAATGAGAGATGAAAAATA 409
Qy 7352 TTTCTACAGAGATATCTTGTGTAAAAAGCGGATGAGAGATCTCAGGAGAGAGAAATTA 7411
Db 410 AATCGAAAACTGAAATAGGTGAGAAATTAACGAAAAAGTATGAAAAAAGTACCTGAG 469
Qy 7412 AAGCATTTGTGAAGACTTGATTTATTTGAAAAAATGATGATTTTACAACAGAGACA 7471
Db 470 AAGTAGCTGAAGAAATTAGTTGAAAAAGTAGAGTGAAGAGTGAAGAAATTTAGTTGAA 529
Qy 7472 AGAATTAATATCACTTCTGCTGCTGTTGGGAATCGACAGTCTTGCTTCCGCATCAGGAA 7531
Db 530 AAGTAGATGAAAAAGTACTGAGAGAGATGATCAAAAAAGTATGAGAGAGATGAACTGAA 589
Qy 7532 CAGTGCAGTTACAAATATTAAGAAAAATCCGAGTTACTGTTGAAAAATCTTTTGTGA 7591
Db 590 AATTAAATGAAAAAGTATGATGAGAAAGTAACTGAAGAAATTAATGAAAAAGTAGATGA 649
Qy 7592 AAGCAGCTGAAAAAGTAAATGTTAGATCGGATATTAACGAGAAATGTTGCTTTAACAGAT 7651
Db 650 AAGTGTCTGAAGAAATTAATGAAAAAGTAGAGTGAAGAAAGTGTGGAAGAAATTAATGAA 709
Qy 7652 ATCAAGTCTCTG--TAGAGACATTGGGAATAGAGCTGCTCATGAGAAATTAATTTCTA 7708
Db 710 AGGTAGCTGATGATTAATTTGAAAAAGTAGAGTGAAGAAAGTGTGAAAGATTTAATTTGAA 769
Qy 7709 ATGGAAGATCAAAATATCAGATTTAAAAATTTCTAGCTATTAGAAAAAATATATGATGTTA 7768
Db 770 AGGTAGCTGATGATTAATTTGAAAAAGTAGAGTGAAGAAATTTAGTTGAAAAAGTAGATGA 829
Qy 7769 TTGTAAAAAGTAAATCGGATTTGAGAGCGGAGCAAAAAGATTAACCTGAGAGCGGTAG 7828
Db 830 AAGTGTGAAAGAAATTAAGTTGAAAAAGTAGAGTGAAGAAAGTGTGAAAGATGATCAAA 889
Qy 7829 CTGCGCGAGCAATTTCTCAAAAGCAAGAAATGAATGAATTCAGAGGTGAAATTTGAGA 7888
Db 890 AAGTAGATGAGAGAAATTAAGTGAAGATTAATTTGAAAAAGTAGATGAGAGAACTGGAAG 949
Qy 7889 AGAGTATTTTCAATGAAGAAAAATAGAGTAATGAGCCCTTTCTAAGAAATTTGAGAGAAA 7948
Db 950 AATTAAATGAAAAAGTAAAGTAGAGTGAAGAGTGTGAGAAATTAATTTGAAAAAGTAGATGA 1009
Qy 7949 TCAATGTCAAAGTGAAGAAAAAGAAACAGAGTACTGCTGATCTCAAGAGCTTCTGTAG 8008
Db 1010 AAGTGTCTGAAGAAATTAATTTGAAAAAGTGAAGTGAATTTAGTTGAAAAAGTAGCTGAAG 1069
Qy 8009 GAGCAGTGAAGGAGGAGAGAAATTAATTTCCGAGAGCAAAAGATCCGGAAGCTCTTATTTGA 8068
Db 1070 AATTAGTTGAAAAAGTAAAGTAGAGTGAAGAGTGAAGAAATTAAGTTGAAAAAGTAGATGA 1129
Qy 8069 AAGTTAGTCAAAATCCGAGAGAGATATTTTTCATGCAATATGATGAAATATGAGAGCAA 8128
Db 1130 AAGTAGCTGAAGAAATTAAGTTGAAAAAGTAGAGTGAAGAAATTAAGTTGAAAAAGTAGCTGA 1189
Qy 8129 CACATTAATGAAGATGACAGAGCTTTCTAAGCAGTAAACAGTTCTGTATTTGGAGAGAG 8188
Db 1190 AAGTAGCTGAAGAGTGAAGTGAAGAGTGAAGAAAGTATACCTGAAGAGTGTGAAGAG 1249
Qy 8189 TTGAGATCAACAAGCAGAGAGTACTGCTGACGAGTAAACCTATGAGAGAGTGTGAAGAG 8248
Db 1250 TAGCTGAAGAGTACTGAGAAAGTGTGAGAGAGTGAAGAGTGAAGAAATTAAGTTGAAAAAG 1309
Qy 8249 GAAATTTGTTCAAGAAC--AAATCGATTTGAATGCAATTTCTAAGTAGAGAGGTTTGATG 8305
Db 1310 TAGATGAAGAGTACTGAGAAAGTGAAGTGTGAAGAGAGAGTGAAGAAAGTGAAGAGAG 1369
Qy 8306 AAGATTAAGTAACTGCTAAATCTTCTGTGTATCAGAGAAATGAGAGAGAAATTTGCCGAG 8365

```

```

Db 1370 TAGTTGAAGAGTGAAGTGAAGTACTGAAAAAAGTAGTGAAGAGAGAGTGAAGAG 1429
Qy 8366 CAGAGTGAATACTTCTTACAGCACAAAGTAACTAGTAATCCGTAGTCTTTAGGAAAGC 8425
Db 1430 TACTTGAAGAGTAAATTAAGAGAGTGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 1489
Qy 8426 AAGATTGAAAAATATGATTTACCAAAAAATATATTTTCAAGAG 8470
Db 1490 TAGTTGAAGAGCAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 1534

RESULT 13
AAA70099
ID AAA70099 standard; DNA; 3579 BP.
XX
AC AAA70099;
XX
DT 07-NOV-2000 (first entry)
XX
DE Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:232.
XX
KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW antimalarial; malaria; protozoacide; infection; insecticide; de.
XX
OS Plasmodium falciparum.
XX
PN WO200025728-A2.
XX
PD 11-MAY-2000.
XX
PF 05-NOV-1999; 99WO-US026796.
XX
PR 05-NOV-1998; 98US-0107131P.
XX
PA (HOFF/) HOFFMAN S.
PA (CARU/) CARUCCI D.
PA (GARD/) GARDNER M.
PA (VENT/) VENTER J C.
XX
PI Hoffman S, Carucci D, Gardner M, Venter JC;
XX
DR WPI; 2000-365347/31.
XX
PT Proteins encoded by chromosome 2 of the human malarial parasite,
XX PT Plasmodium falciparum, useful as anti-malarial vaccines and in the
XX PT diagnosis of P.falciparum infection.
XX
PS Disclosure: Page 457-458; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II). (I) and
CC (II) are useful for the development of vaccines against P. falciparum
CC infection. (I) and polyclonal antisera or a monoclonal antibody raised to
CC immunogens comprising the sequences of (I), are useful in the detection
CC of infection with P. falciparum. Furthermore, (I) (especially when they
CC are rifins or secreted or membrane proteins) can aid the identification
CC of drugs to treat or prevent P. falciparum infection, or they can be used
CC to identify drug resistance in P. falciparum. Sequencing of the
CC Plasmodium chromosome 2 and the subsequent identification of proteins
CC encoded by it will help to expand our understanding of parasite biology,
CC a process hampered by the complexity of the parasitic lifecycle, and
CC provide new targets for vaccine and drug development. Parasite resistance
CC to drugs and mosquito resistance to insecticides have led to a resurgence
CC of malaria in many parts of the world, and there is a pressing need for
CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
CC represent nucleotide and protein sequences given in the present
CC invention, but which are not specifically mentioned within the
XX specification
XX
Sequence 3579 BP; 1904 A; 398 C; 552 G; 725 T; 0 U; 0 Other;

```

Query Match	0.9%; Score 87.4; DB 3; Length 3579;
Best Local S.Milarity	42.7%; Pred. No. 9e-07;
Matches	616; Conservative 0; Mismatches 821; Indels 6; Gaps 3;
OY	7110 AGAAAAAGTAAATGTAACGGCAAAAGAGAAAAAATTAAGCAACAGCAGCAATATGC 7169
DB	1362 AGAATCAGAAAAAGATTAAGAAAGATTCAGAAAAAGACAAAGAAAAAATCGAAGAAAGATGA 1421
OY	7170 AGGAATCGAGAGAGAGCAATCGGAGCCAAATCTTGGTAAATATTTTGGAAAGCTGT 7229
DB	1422 AGAAAAAATCTGAAGCGAAAAAGAACGAAAGTATACAAAAAGAAACAGATGTAGATGA 1481
OY	7230 AGAAGATGAAAAAATTTCTGAAGAAAAAGAAACAGAAATTTTAAAACTTTAGACGAAGT 7289
DB	1482 AAAAAAAGAAAAAGGAGAAATATGAGAGGGAACAGATATGAGAAAGACAAAGAAAAAGAA 1541
OY	7290 TAAACAAGAACAGATTAATAAAAGTAATGATGCTACGAAAAAATCTTACATCGACAG 7349
DB	1542 AGAAGACGACGAAAGAAACAAAAGTAAAGAAAAAGAAAAAGACGAAAGAGGAAAC 1601
OY	7350 TATTTCTACAGAGACTTCTGTGTAAGCCGATAGAGAGATCTCAGGAGAAAGAT 7409
DB	1602 AGATTATAGAGAGATACAGATGATTCAGAACAAAGATGAAAGAAACAAAAGTAAAGAAAA 1661
OY	7410 TAAAGCCATTGGAAGACTTCTGATATTATGAAAAAATGTAGATTTACACAGAGGA 7469
DB	1662 GAAAAACAGAAAGAGACGAAAGAAACCTGAAGAAACGAAAAAGAAACAGAAAGTAAAGAAA 1721
OY	7470 CAAGATTAATATCACTTCTACTGTGTGGAACTGCAAGCTCTTCCGATCAGG 7529
DB	1722 AAAGAAAAACAGAAAAAGACGAAAGAGGAAACAGATTATGAAAGATATACAGATGATTCAGA 1781
OY	7530 AACGTGCGAGTTACAAATATTTAA--AGAAATTCGCGAGTACTGTTGAAAAATTTCTT 7586
DB	1782 CAAGATATGAGAAACAGAAATGAGAAACACAGCCGACGACGAAAGAAAGAAAGAA 1841
OY	7587 TGTGAAACAGCTGAAAAAATGTAATGTTAGATCGATATTAACAGGAAATGTGCTTTAAC 7646
DB	1842 GGGAAACAGATGATGAGAAAGAACAAAGTAAAGAAACACACCTAGACGACAAAGAAAGAA 1901
OY	7647 AGCATATCAAGTCTCTGTAGAGCATTTGGAAATAGAGCTGCTATGCAAGATTTAAATTC 7706
DB	1902 CGGAGAAAGAGATTAAGAAAGACGAAAGAAAAAAGACAAAGAAAGATGACAAAGAAAGATGA 1961
OY	7707 TAAATGAGATCAATATCAAGTATTAATAATTTCTAGCTATTTAGAAAAAATATTGATGT 7766
DB	1962 CAAGAAAAAAGACAAAGAGATGACAAAGAAAAATACAAAGAAAGATGACAAAGAAAGATGA 2021
OY	7767 TATTGTAAGAATTAATCGAAATTTGAGACGGAACAAAGATTTAACCGTAGAGCGGT 7826
DB	2022 CAAGAGATGATGACAAAGAAAGAAAGACAAAGAAAGATTAACAAAGAAAGAAAGATTA 2081
OY	7827 AGCTGCCGAGCCATTAATCTCAAAAGCAAGAAATGAATGAAGTTGAGAGTTGAAATTGA 7886
DB	2082 CAAGAAAAAAGACAAAGAAAGATGACAAAGAAAAAAGACAAAGAAAGATGACAAAGAAAGATGA 2141
OY	7887 GAAAGTATTTTCAATGAAGAAATAGATTAAGCTTAAAGCAATTTGAAAGAA 7946
DB	2142 CAAGAGATTAACAAAGAAAAAGACAAAGAAAGATTAACAAAGAAAAAGCAAGAAAGATGA 2201
OY	7947 AATCAATGTCAAAGTGAAGAAAAAGAAACAGAGTGACTGCTGAATCTCAAG--GAGCTTCT 8004
DB	2202 CAAGAAAAAAGCTGTAAACATGTAAAGAAATTAAGAAAGAAAGAAAGATGATGACTA 2261
OY	8005 GTAGAGCAGTAGACAGGGCAGAGAAATTAATTTCCAGACAAAGAAAGTCCGGAAGCTTTAT 8064
DB	2262 TGATGAAAGTTTGAAGAAAGAAATTAATTCACATTAATATAGCAATTTGGACAAACACA 2321
OY	8065 TTGAAAGTTAGTACAAATCCGAAAGAAATTTTTCATGCGATTAATGTGAATGTGAA 8124
DB	2322 ATATTTCCTACTATTAATCCATTTAGAACAAAGAAATTAATCAACTATTAATTAATA 2381

OY	8125 GCAACACATTAATAATGAAGATTAACAGCAGTTTCTAAAGCAGTAAAGGTTCTGTATTGGGA 8184
DB	2382 ACAACAGAAAAATTAATCAACCCACATCATTTATTAATAACACAGCAAAATCATTAATCCACA 2441
OY	8185 GGAAGTTGAGTCCACCAAGGCAAGAACTACTGCTCAGGTAAACTATGTGTGAAGTTGAG 8244
DB	2442 TCATATTTTACAGAGCAAGAAAAACATCACCCACAGGTATATCCAAAGAAACACATTA 2501
OY	8245 GAAGAAATTTGTTCAAGAACAAATCGATGAAATGCAATTTCTAAAGTAAAGGTT--TGGG 8303
DB	2502 TAAATTAATGTTCTTATATCTTAAAAAAGGGCTTGAAACCCAAACTCATACATGTAA 2561
OY	8304 TGAAGATTAAGTAACTGCTAAATCTTCTGTAGTACAGAAATGAGAGGAAATTTGCCG 8363
DB	2562 AGAAGATCAACCTTAATTAATAACAGGGCTGTGAAGGACAAAGAACACATGTGTGATGA 2621
OY	8364 AGCAGAGTGAATCTTCTACAGCAACAAAGTAATCTGAATTCGGATGTTGTTACGAA 8423
DB	2622 TATGCATTAACAAATCAAAAGAACATTAAGAAATTTAAATAACAAACCGATGTAAACAA 2681
OY	8424 GCAAGATTATGAATAATATGATTAACAAAAAATATTTTCAAGAGCAATGCTCTGC 8483
DB	2682 AGCAAGTCATATATATTAATTAATTCATCAGAAAAACAAATTAACATGTATTAATAGTC 2741
OY	8484 TTTAAATGATCAAAAGATGAAGCAATATAGAAATCTTTAGCGGTAGCCGTTGTCATGC 8543
DB	2742 TCTGAAAAAACAATTTGAACATGTATTAATTAATAGCTCTCTGAAAAACAATTTGAACATGT 2801
OY	8544 ACA 8546
DB	2802 ATA 2804

## RESULT 14

AA087587 ID AA087587 standard; DNA; 1686 BP.

AA087587;

DT 27-AUG-2003 (revised)

DT 19-DEC-1995 (first entry)

DE DNA encoding Leucocytozoan protozoa structural protein epitope.

KW Leucocytozoan protozoa; structural protein; epitope; vaccine; fowl; leucocytozoanosis; treatment; ss.

OS Leucocytozoan.

PN JP0708995-A.

PD 04-APR-1995.

PF 10-SEP-1993; 93JP-00226078.

PR 10-SEP-1993; 93JP-00226078.

PA (DOBU-) DOBUTSUVO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.

PA (NISS-) NISSEIKEN KK.

DR WPI; 1995-167252/22.

DR P-PSDB; AAR70491.

PT Immune inducing polypeptide against Leucocytozoan protozoa - useful in production of vaccines for treatment of Leucocytozoanosis in fowl.

PS Claim 1; Page 12-14; 20pp; Japanese.

CC AA087587-89 encode polypeptides having a whole or partial epitope of a structural protein of Leucocytozoan protozoa (see AAR70491-93). The polypeptides and DNA encoding them are useful in the production of vaccines for the treatment of Leucocytozoanosis of fowl. (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 U; 0 Other;

Query Match 0.8%; Score 81.8; DB 2; Length 1686;  
 Best Local Similarity 42.4%; Pred. No. 8.8e-06;  
 Matches 513; Conservative 0; Mismatches 692; Indels 6; Gaps 1;

QY 1112 AAAAAGTAAATGTAACGGCAAAAGGAAAGAAAAATTTAAGCAACACGCAAAATGCAAG 7171  
 DB 302 AAGATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 361  
 QY 7112 GAATGAGGAGGACCAATCGGAGCCAAATGCTTGTAAATATTTTGAACAGCTGAG 7231  
 DB 362 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 421  
 QY 7232 AAGATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 7291  
 DB 422 AAGATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 481  
 QY 7292 ACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 7351  
 DB 482 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 541  
 QY 7352 TTTCTACAGAAAGATCTCTGTAAAGGAGGAGAGATCTGAGGAGAAAGAAATTA 7411  
 DB 542 AAGATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 601  
 QY 7412 AAGCATTTGTGAAGCTTCTGTATTTATTTGAAGAAAGTATTAACAAGAGACA 7471  
 DB 602 AAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 661  
 QY 7472 AGAATTAATATCATCTTCTACTGCTGTGGGAACTGACAGCTTGTGCTCCGATACGAA 7531  
 DB 662 TATATGTAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 721  
 QY 7532 CAGTGGCAGTTCAATATTTAAAGAAATTCGGAAGTTACGTGTGAAATTCCTTTTGA 7591  
 DB 722 AAGTGAAGAAAGTAAACCTTAAGATGATTAAGAAAGAAAGAAAGAAAGAAAGAAAG 781  
 QY 7592 AAGCAGCTGAAAGAAAGTAAATGTTAGATCGGATATTAACAGAAATGTTGCTTTAACAGCAT 7651  
 DB 782 AAGAAAGAAAGAAAGTAAAGTAAAGAAAGAAAGTAAAGAAAGTAAAGTAAAGTAAAG 841  
 QY 7652 ATCAAGGTCCTGTAGAGCAATGGGAAATAGAGCTGCTTAAGCAGAAATTTCTAATG 7711  
 DB 842 AAGTAAAGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 901  
 QY 7712 GAAGATCAAAATATCAATTTAAAGAAATTTCTAAGCTATTTAGAAAGAAATTTAGTGTATG 7771  
 DB 902 AAGATCCAAATTTGTCATCAATATTAAGATATTAAGTAAATTAAGAAAGAAAGAAAG 961  
 QY 7772 TAAAGATTAATCGGAATTTGAGAGCGGAAAGAAAGTAAACCGTGAAGCGGTAGCTG 7831  
 DB 962 AAATCTAATCAGCAAAATTTAATCAATTTAAAGTAACTAATCAAAATCCGAAATTAAT 1021  
 QY 7832 CCGAGCCATTAATCTCAAAAGCAAGAAATGAATTCAGAGGTTGAATTTAGAGAGA 7891  
 DB 1022 CAACATCAATATTTATTTGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1081  
 QY 7892 GATATTTCAATGAAGAAATTAAGTAACTAGAGCTTTCTAAGAAATTTGAAGAGAAATCA 7951  
 DB 1082 ATGAAGCTACTAGAGCCGGAATTCGAGCACTAGTTGAGAAAGTGTGGAAGAAACAC 1141  
 QY 7952 ATGTCAAGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 8011  
 DB 1142 CAGTATTTGAAGAAAGAAAGTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1201  
 QY 8012 CA-----GTAGCAGGAGCAGAAATTAATTTCCAGACAAAGATGCGGAAAGCTTTAAT 8065  
 DB 1202 AAACACCTGTAGTTGAAGAAAGAAAGTATGAAGAAAGAAAGTATGAGAAAGATGTGCG 1261  
 QY 8066 TGAAGATTAGTACAAAAATCCGGAAGAAAGTATTTTTCATGAGATTAATGGAATATGGAAG 8125

DB 1262 TCGAAGAAACACCAATAGTGAAGAAAGATGATAGAAAGAAAGTAAATGATGGAAG 1321  
 QY 8126 CAACATTAATTAATGAAGTAAAGCAAGTTTAAAGCAGTAAAGTTCGTATTTGGAG 8185  
 DB 1322 TAGTGAAGAAAGAAAGCAAGTATTAAGAAAGAAAGTAAAGAAAGTCTCTTATGTTGAAG 1381  
 QY 8186 GAGTTGAGTCAACCAAGCAGAAAGCTACTGCTGAGGTAAGTAAATGATGTAAGTTGAGG 8245  
 DB 1382 AAGTATGTAAGAAAGAAAGCAACCAATAGTGAAGAAAGTGTGAAAGAAAGCAACCAAGTATG 1441  
 QY 8246 AAGAAATTTGTTGAGAAACAAATGCATTAATTCATTAAGTAAAGTGTGATG 8305  
 DB 1442 AAGAAAGAAATGTTAGAAAGAAAGCAACCAATAGTGAAGAAAGATGATGAAAGAAAGCAAC 1501  
 QY 8306 AAGATTAAGTA 8316  
 DB 1502 TAGTTGAAGAA 1512

RESULT 15  
 ADP85917  
 ID ADP85917 standard; DNA; 1300 BP.  
 XX  
 AC ADP85917;  
 XX  
 DT 26-AUG-2004 (first entry)  
 XX  
 DE Synthetic construct #1.  
 XX  
 KW Nanopore data analysis; polymer; polynucleotide;  
 KW single-molecule resolution; de.  
 XX  
 OS Synthetic.  
 OS  
 PN US2004110205-A1.  
 PN  
 PD 10-JUN-2004.  
 PD  
 PF 23-SEP-2003; 2003US-00668749.  
 PF  
 PR 23-SEP-2002; 2002US-0412959P.  
 PR  
 PA (WANG/) WANG H.  
 PA  
 PI Wang H;  
 PI  
 DR WPI; 2004-467656/44.  
 DR  
 XX  
 PT Nanopore analysis system for analyzing polynucleotides includes nanopore  
 PT device having structure with aperture, and nanopore data analysis system  
 PT that can generate nanopore data points corresponding to each target and  
 PT non-target polymer.  
 PT  
 PS Disclosure; SEQ ID NO 1; 17bp; English.  
 PS

The invention relates to a nanopore analysis system that comprises a  
 CC nanopore device and a nanopore data analysis system. The nanopore device  
 CC has a structure having an aperture. The nanopore data analysis system can  
 CC generate nanopore data points corresponding to each target polymer and  
 CC each non-target polymer traversing the aperture; form a distribution  
 CC pattern of the data points; and analyze a distribution of target polymer  
 CC data points in the distribution pattern. The distribution pattern  
 CC includes data cluster(s). The nanopore data analysis system is stored on  
 CC a computer-readable medium. A mechanism is provided for analyzing the  
 CC distribution of target polynucleotide data points in the distribution  
 CC pattern. The nanopore data analysis system can also analyze the  
 CC distribution of the non-target polynucleotide data points. It can also  
 CC determine a ratio between the target polynucleotide data points and the  
 CC non-target polynucleotide data points. It can further analyze the  
 CC distribution of target polynucleotide data points between the two data  
 CC clusters; compare the distribution of the target polynucleotide data  
 CC points between the two data clusters to a phosphorylation state standard  
 CC distribution; and determine a ratio of phosphorylated target

CC polynucleotide to non-phosphorylated target polynucleotides. It can also  
 CC determine a cluster score for the target polynucleotide data points in a  
 CC defined area; and compare the cluster score for the target polynucleotide  
 CC data points to a cluster score for a chemical integrity standard density  
 CC distribution for the defined area in a distribution of a target  
 CC polynucleotide standard. The invention is useful for analysing polymers,  
 CC such as polynucleotides. The inventive nanopore analysis system  
 CC potentially provides high speed sampling with single-molecule resolution,  
 CC which may enable unprecedented dynamic range and sensitivity in analysis  
 CC of samples containing charged polymers such as polynucleotides or  
 CC polypeptides. The present sequence is a synthetic construct used in the  
 CC invention.

XX Sequence 1300 BP; 1300 A; 0 C; 0 G; 0 T; 0 U; 0 Other;

Query Match 0.8%; Score 80.4; DB 12; Length 1300;

Beet Local Similarity 41.4%; Pred. No. 1.5e-05;

Matches 537; Conservative 0; Mismatches 761; Indels 0; Gaps 0;

QY 8426 AAGATTATGAAATATGATTCACAAAAAATATATTTTCAGAACTCATGCTCTTGCTT 8485  
 DB 1 AAA 60  
 QY 8486 TAAATGATACAAAGATGAGCAATATGAAATCTTTAGCGGTAGCCGGTGCATGCAC 8545  
 DB 61 AAA 120  
 QY 8546 AAGAAACAACAAACATTTACAGATCAACAAGTTAACTTCAACTGTAAATGAGAG 8605  
 DB 121 AAA 180  
 QY 8606 GAAAGTATCTCACTTCGTGCAGAAAGCTTTGGCTAAATAATGAAATTAAGAAATGTA 8665  
 DB 181 AAA 240  
 QY 8666 AAGGAACCTGAGAGAGCTTATGTCGAGCGGAAACGACCGCTGAAAAATTAACAAGA 8725  
 DB 241 AAA 300  
 QY 8726 GATCTACAGAGACATTTGGTTCAGAGAAATTGGGAATTTGAGATTAATTAAGATTTG 8785  
 DB 301 AAA 360  
 QY 8786 CAAGAGATATATCGAATTTGTAAGATCAACGAGACGAAACAAGAGAGCTTTGTCGAA 8845  
 DB 361 AAA 420  
 QY 8846 AGAATGATTTCTGTGAAAAATACAAATTTCAAGGGGAAACAAATATCATTCATTGAAGTA 8905  
 DB 421 AAA 480  
 QY 8906 AAGCGAATTTGTTGGAACCGGAAGTGTAAATGTGATGCTTTGAATGAATCTTGATGTAG 8965  
 DB 481 AAA 540  
 QY 8966 ATCTCAAGAAAAAAGTGTGCTATGTGTAATTTGTTAATGAAATGTTGATGTAATA 9025  
 DB 541 AAA 600  
 QY 9026 ATGTGATTTAGAAAAATGTGAGAGCCAAATCGAAGACATGTAATTTGAGAACTACTG 9085  
 DB 601 AAA 660  
 QY 9086 GAAAAACAAGATATCAAGATTTTACAGAGCAAAAGTAATTTCTTGGAAGAAAGAGAGC 9145  
 DB 661 AAA 720  
 QY 9146 CTGACAGCTGACGCTGCAATATGCAATGTACATTTCAATGAGATGATTAATAATT 9205  
 DB 721 AAA 780  
 QY 9206 TGGCAAGCAGTATGCTCTCTCAATTAATACCAAAAAATTCAAAAATAATTAATCTT 9265

DB 781 AAA 840  
 QY 9266 TAGCATCAAGTATGATTCGAATGTAATGTTTCATGGGGTGTGAAAGCAAGAGTGCAG 9325  
 DB 841 AAA 900  
 QY 9326 GAGCCAAAGCAGATGATGTAAGAAATCAATTAATAGAACTAATTAATGATTTAG 9385  
 DB 901 AAA 960  
 QY 9386 CAGAAAAAATTTAAACAGAGGAAACATCAATGTAATGTCGAGATGATTAATAATA 9445  
 DB 961 AAA 1020  
 QY 9446 ATATAGTATGACAAATTTCTAAGGCTATTGCGAGATGCCAAAGTCATGCTGCTGCTT 9505  
 DB 1021 AAA 1080  
 QY 9506 CGGCAACTGCCCATTTGTAAGAAAAATGAAAGTAAATTTAATGCGATCCGAGAAATTTA 9565  
 DB 1081 AAA 1140  
 QY 9566 AAAATTAATCTGCAAGATTTGGAAGGAAAGCTAATTAATAAAGCTGGTACATCTAATC 9625  
 DB 1141 AAA 1200  
 QY 9626 AGTATGACTGGTATACGATTAATATATATACATGCACTTCTTGAAAAAGCATACAAAAAT 9685  
 DB 1201 AAA 1260  
 QY 9686 TGACATATCAATCAAGAGAGAGAGAAAAAGGAAAAAA 9723  
 DB 1261 AAA 1298

RESULT 16  
 ADA89806/C  
 ID ADA89806 standard; DNA; 4965 BP.  
 XX  
 XX ADA89806;  
 XX  
 XX 20-NOV-2003 (first entry)  
 XX  
 DE Staphylococcus aureus antigenic partial gene sequence locus 35B.  
 XX  
 XX anti-genic protein; vaccine; Staphylococcus aureus; pathogenic organism;  
 XX antibacterial; neuroprotective; immunosuppressive; anti-inflammatory;  
 XX antitumor; immunostimulant; ophthalmological; pathogenic microbe;  
 XX bacteremia; septic shock; organ infection; skin infection;  
 XX bacterial basal colonization; bacterial eye infection; septicemia;  
 XX tuberculosis; food poisoning; blood infection; peritonitis; endocarditis;  
 XX sepsis; meningitis; pneumonia; stomach ulcer; gonorrhea; toxic shock;  
 XX necrotizing fasciitis; impetigo; histoplasmosis; Lyme disease;  
 XX gastro-enteritis; dysentery; shigellosis; skin disorder; gene; ds.  
 XX  
 OS Staphylococcus aureus.  
 XX  
 PN WO2003011899-A2.  
 XX  
 PD 13-FEB-2003.  
 XX  
 PF 02-AUG-2002; 2002WO-GB003606.  
 XX  
 PR 02-AUG-2001; 2001GB-00018825.  
 PR 09-JAN-2002; 2002GB-00000349.  
 XX  
 PA (UTSH-) UNIV SHEFFIELD.  
 PA (BIOS-) BIOSYNEXUS INC.  
 PI Foster S, Mond J, Clarke S, McDowell P, Brumel K,  
 XX  
 DR WPI; 2003-256434/25.  
 XX

PT New antigenic polypeptides from *Staphylococcus aureus* or *S. epidermidis*,  
PT useful as a vaccine for immunizing humans against e.g. bacteremia, septic  
PT shock, septicemia, tuberculosis, meningitis, pneumonia, gonorrhea or  
PT impetigo.

XX Claim 1; Page 68-69; 1899p; English.

CC The present invention describes an antigenic protein or its part, which  
CC is for use as a vaccine. The antigenic protein is encoded by an isolated  
CC DNA molecule of *Staphylococcus aureus* or *S. epidermidis* partial gene  
CC sequences (designated dasa and dna SE, respectively) and which encodes a  
CC protein expressed by a pathogenic organism. Also described: (1) a vaccine  
CC composition comprising at least one antigenic protein; (2) a method of  
CC immunizing an animal against a disease or condition caused by a  
CC pathogenic microbe by administering the antigenic protein or the vaccine;  
CC (3) an antibody or its binding part obtainable by the method above; (4)  
CC preparing a hybridoma cell line producing monoclonal antibodies; (5) a  
CC hybridoma cell line produced by the method of (4); and (6) identifying  
CC opsonic antigens expressed by a pathogenic microbe. The antigenic  
CC proteins have antibacterial, neuroprotective, immunosuppressive,  
CC antiinflammatory, antitumor, immunostimulant and ophthalmological  
CC activities, and can be used in vaccines. The antigenic proteins or  
CC vaccines can be used for immunising an animal (specifically a human)  
CC against a disease or condition caused by a pathogenic microbe, e.g.  
CC bacteraemia, septic shock, organ infection, skin infection, bacterial  
CC basal colonisation, bacterial eye infections, septicemia, tuberculosis,  
CC bacteraemia-associated food poisoning, blood infections, peritonitis,  
CC endocarditis, sepsis, meningitis, pneumonia, stomach ulcers, gonorrhoea,  
CC strep throat, streptococcal-associated toxic shock, necrotising  
CC fasciitis, impetigo, histoplasmosis, Lyme disease, gastro-enteritis,  
CC dysentery, shigellosis, *S. aureus*-associated septicemia, food-poisoning,  
CC skin disorders, *S. epidermidis*-associated septicemia, peritonitis or  
CC endocarditis. The present sequence represents a *S. aureus* partial gene  
XX sequence from the present invention.

XX Sequence 465 BP; 1119 A; 870 C; 861 G; 2115 T; 0 U; 0 Other;

XX Query Match 0.8%; Score 79.2; DB 8; Length 4965;

XX Best Local Similarity 41.3%; Pred. No. 3.9e-05;

XX Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

QY 3782 TTCAAGTAAACCAAGTCTTTGTTAAATAGTATTTGGAATGCGAATTAATT 3841  
DB 4395 TTAAATTAACCAACATGCACTGATGAATCAACAGCAATTTTAATTAACAAATTGCAAG 4336  
QY 3842 AAGTAATGCTTTAGTGAAGAACTCAAGTAGCAGACAGAGAGGTTTGAAGCAGTTA 3901  
DB 4335 AAGATTCAGCTACGCCAGATGCAACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4276  
QY 3902 AAGAAAGTGAAGACAAAGAGAAAGTATCTATTTGGAACTTCTGCTTCTATCACTTAA 3961  
DB 4275 ATACTGAATGATGTAAGCAATTCAGCACTTTCAGCAGCACTATTAACGCAAGTTG 4216  
QY 3962 TGAACATGAAGTTTTCGCAAAATCGAAATTAATCACTAGAGAGAGATCTGAAGCC 4021  
DB 4215 ATGAAGCTAAAGCAATGCAAGAGCAGCACTTAATCGGTAAACCAAAAGTTGTGAAG 4156  
QY 4022 AAAAAATGATGTTGATGCTCACTGCTTATCAAGCGGACCCAAAGTAGCAGAGCTTTAA 4081  
DB 4155 AACAGCGGCTAAAGATGAATGATGATTAATCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4096  
QY 4082 ATTCAAGCTGAAGAGTCAATGAACTGTAGGGGCTACTGTGACTGTTGCCAAATTA 4141  
DB 4095 ATGATCAGAACGCTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4036  
QY 4142 ACAACAAATTAATGCTTCTATTAAGTGTGGAGATTAATCACTTAATCGAGCGGAGC 4201  
DB 4035 TTTCAGAGCGGAG 3976  
QY 4202 CAAGAGCTTTTGAAGCACTCAAGTAGCTGCTGAGTAGAGAGAGAGAGAGAGAGAGAG 4261  
DB 3975 AAGACGCTGAAGAGAGATTAATCAATTAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3916

QY 4262 GTTCTGAGCGGAGTTAGAAATTAATCAAGGCGCTGTTCTGTCAATTAAGATGACATG 4321  
DB 3915 CTAATAATGATGTTGATCAAGCTGTGACACTCAAAATCAAGCAATGATTAATCAACTG 3856  
QY 4322 ACGTGAAGCTAGCGGTTGATTAATCTTCATCGAAGAGCTTAATGAATCAATGCTATTG 4381  
DB 3855 GTGCTACACATGAG 3796  
QY 4382 CCAAGATGTCAGAAAG 4441  
DB 3795 GGTATCAAGATATCTTTAAATGCAAAACACTTAATGATGTTTACGCAATTAAGATCAAG 3736  
QY 4442 AAGATTAATAATTTAGAAATGATCGGTATTAATACACTGAGAAATGTTATTAATACGA 4501  
DB 3735 CAGTTCGATATTCAGAGTATTAATGAGATTAACAACTTAAGATTTG-----CGA 3682  
QY 4502 AAGAACATGAG 4561  
DB 3681 AAGATGATGAG 3622  
QY 4562 CGGTTGCTGAGAGAGATTAATCCCTGAGAGAGATGATTAATGCACTCAATCTGTTAAA 4621  
DB 3621 CGACTACTGAG 3562  
QY 4622 ATAAATTAAGCAGAAATGAGTGAAGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 4681  
DB 3561 ATCAAAATTAATTAAG 3502  
QY 4682 CGAAACATGTAATGAG 4741  
DB 3501 TTCAAGCAATGAG 3443  
QY 4742 TTGCTATGAG 4801  
DB 3442 TTGCTATGAG 3389  
QY 4802 ATGACAGATGAG 4861  
DB 3388 -----ACTATGAG 3334  
QY 4862 CAATTAATTCATCTTGGGAGATGTTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4921  
DB 3333 GTTAAATTAATTAATGAG 3274  
QY 4922 CGTGAAGAGCTGTTTGGCAATTAATCTTCAATTAATTAATCTGCTTTGATTAACG 4981  
DB 3273 CAGTCAAAAGAGTCAACCACTTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3214  
QY 4982 GAAAGAGTAAATCTTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5041  
DB 3213 AATTAATCAAGCTGAGCTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3154  
QY 5042 ATTCTCATTTTCAACAGCTTTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5101  
DB 3153 AACAGAAATTAATGATGAG 3094  
QY 5102 GAGAGATGATCTGCTCAATCGTGTCTGATGAAGAGAGAGAGAGAGAGAGAGAGAG 5161  
DB 3093 ATGTCATCAATCAATCAACAAATGAATATGTTGATTAATGCAATTAAGAGAGAGAG 3034  
QY 5162 AGTTTGAAGAGATTAAGTTCTTCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5221  
DB 3033 AATTAATCAAGTAAACATTTAGTGAATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2974  
QY 5222 TTGCGGAATGCAATGAG 5281  
DB 2973 ATGCAATTAATGCTAAAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2914  
QY 5282 ATATTGAG 5341  
DB 2913 TTGCTGAAG 2854  
QY 5342 ATCAAGATGAG 5401

```

Db      2853 AAGTACTCTTAAAGATGACATTTGAAGTTCAATTCATATGACTTAGATTAATTAACG 2794
Qy      5402 CAGTGGAGTTGGAGAGCAAAAGAGCCTGTGTGCAAGAGCTCTGCAAGTACTACT 5461
Db      2793 ATTACACATTCACACAGGTAAAGAAAGATCAGTACACAAATTTATATGTGTTAGAG 2734
Qy      5462 TGAATTAAGCAGTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTAGAGAG 5521
Db      2733 ATCAGAAAGAAAATATATTTTCAAGCTGACACTTAATGCAACAGAAATGAAAAGCAAG 2674
Qy      5522 AAAATTAATGAAATTAAGAAAAGCAAAATGTTAATGTTCTAGCTGAAAATACAGTCAAG 5581
Db      2673 CAATTAGCAGTTGACCAAAATG---TTCAAACCTGCATTAGAAAAGCATTTAATATGTTG 2617
Qy      5582 TGGTCAAAATGCGACAGTCTTTCGCGAGCAAGTGGACAAAGCTCAGTAGAGAGCGAG 5641
Db      2616 TGGATTAATGTGACGTTGATGATGATTAACAAAGTAAAGCAGCAATGATGTATTC 2557
Qy      5642 TAGCAGTTAATTAATTAACACAAATTAATCTTGCACATTAATAAATAGTACTCAAAATG 5701
Db      2556 AAGTAGATGCTACTGTTAACTTAAGCAACCAAGCTATTGAAGTTAAAGCAAGATTA 2497
Qy      5702 TAGCAATGCTTTGTTAAAG-----CAAATCTCATTCATCTTA 5740
Db      2496 CGAAAGAAATCTATTGATCAAAAGTACCAAGTTAATGCTGAAGAAAACCTGAAGCATTTAG 2437
Qy      5741 TTTAAACAATTGGAGCTGGAGCTGGAGCTGGAGAGCTGGAGAGTGGACAGTTCTG 5800
Db      2436 CATATATTAAACAAATTAACAGATCAAGCTAAACAAAGTATTACTGATGCAACACAACTG 2377
Qy      5801 TAGCAGTAATTAAGATTGTAATTAATAGCATAGCAGAAATTAATCATGCAAAATCAGCTG 5860
Db      2376 CTGAAGTTGAAAAGCAAGCAAGCTCAAGGACTTGAAGCAATTGATTAACATCAATGCACT 2317
Qy      5861 CGAAGGAAATGTGCGAAGTTATTACAGAGTCTGATGCGGTAAATTCCTAATTATGCAAGAA 5920
Db      2316 CAACGAAAACAAAGAACTATCGAAGATTGAAGAACTGCACTAGACCAAGATTGAAGCAG 2257
Qy      5921 CAGTGTGAGAGTGGCCGCTGACGCAATAGAGCCTCAACAGTGTGAATGAATTAAG 5980
Db      2256 GTGTAATAGTCAACGCTGATGTCTACACAGTAAGAAAAGAACGCTTTTCAAGATGCTTTAG 2197
Qy      5981 GATCTACAAAGCATATATGTAATTAATAGTATTCATAGATTCCTAAAGAAACAGATGAT 6040
Db      2196 AAGACATTTTATCAAAAGCACTGAAGATATTCTGATCAACTAACAATGAGAAATCG 2137
Qy      6041 ATATTACTACTCAAGGCAAGTAGATTAAGTGTGATTAAGTATTCAAAAATCTTAATA 6100
Db      2136 CTACTGTCAAAAATAGTGGCTTGAACAACCTTAAGCAACAGTATTATCTGAAAGTTA 2077
Qy      6101 TTAAGAGAGCTTATCACAAAAAGAAAATAAGTAAATTAAGATTTGTTACCAATA 6160
Db      2076 AGAAAAATGCTTTGGAAGCAATCAGAGAAAGTGTGTTAACAGCAATAGAAAATTAATAA 2017
Qy      6161 GTTCAGTACTACTACTT 6178
Db      2016 ATGCAGATGCAAGTGCAT 1999

```

RESULT 17  
AAS52179  
ID AAS52179 standard; DNA; 7434 BP.

AC AAS52179;  
DT 13-FEB-2002 (first entry)  
XX  
DE Staphylococcus aureus DNA for cellular proliferation protein #596.  
XX  
KM Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;  
XX antibacterial; drug design.

```

OS      Staphylococcus aureus.
XX
PN      WO200170955-A2.
XX
PD      27-SEP-2001.
XX
PF      21-MAR-2001; 2001WO-US009180.
XX
PR      21-MAR-2000; 2000US-0191078P.
PR      23-MAY-2000; 2000US-0206848P.
PR      26-MAY-2000; 2000US-0207727P.
PR      23-OCT-2000; 2000US-0242578P.
PR      27-NOV-2000; 2000US-0253625P.
PR      22-DEC-2000; 2000US-0257931P.
PR      16-FEB-2001; 2001US-0259308P.
XX
PA      (ELIT-) ELITRA PHARM INC.
XX
PI      Haeelbeck R, Ohlsen KU, Zyckind JW, Wall D, Trawick JD, Carr GJ,
PI      Yamamoto RT, Xu HH;
XX
DR      WPI; 2001-611495/70.
DR      P-PSDB; AAU34320.
XX
PT      New polynucleotides for the identification and development of
PT      antibiotics, comprise sequences of antisense nucleic acids.
XX
PS      Claim 27; SEQ ID NO 4761; 511pp; English.
XX
CC      The invention relates to antisense inhibitors of gene essential to
CC      prokaryotic cellular proliferation, their use in identifying the genes,
CC      their use in the discovery of novel antibiotics, the essential genes
CC      themselves and the encoded proteins. The prokaryotes used are Bacillus
CC      coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC      Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC      useful for the identification of potential new targets for antibiotic
CC      development. The antisense nucleic acids can also be used to identify
CC      proteins used in proliferation, to express these proteins, and to obtain
CC      antibodies capable of binding to the expressed proteins. The proteins can
CC      be used to screen compounds in rational drug discovery programmes. The
CC      antisense nucleic acid sequence is also useful to screen for homologous
CC      nucleic acids which are required for cell proliferation in a wide variety
CC      of organisms. The present sequence encodes an essential prokaryotic
CC      cellular proliferation protein. Note: The sequence data for this patent
CC      did not form part of the printed specification, but was obtained in
CC      electronic format directly from WIPO at
XX      ftp.wipo.int/pub/published_pcc_sequences
XX
SQ      Sequence 7434 BP; 3197 A; 1326 C; 1298 G; 1613 T; 0 U; 0 Other;
XX
Query Match      0.8%; Score 79.2; DB 4; Length 7434;
Beet Local Similarity 41.3%; Pred. No. 4.4e-05;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

```

4082 ATTTACAACTGGAAGTCAATGGAACCTGTAGGGGCTACTGTGACTGTGCCAATTAA 4141  
4083 |||||  
4084 |||||  
3959 ATGATCAGAACCTCTCAACAGAAAGAAAAGAGCTTATCAACATTTAGCAACGCGAG 4018  
4142 ACAACAAAGTAATGCTTATTAAGTGGGAGATATCTAATACGTTAATCGAGCGGAG 4201  
4019 TTACAGACCGGAAAATTAATTTACAGCTGCAACTGATGATTAATGTGTAGATCAGGCGA 4078  
4202 CAAAAGCTCTTTAGCAACCACTCAAGTACTGCTGCAAGTGAAGAGGAGGAGCAATTA 4261  
4079 AAGAGCTGGAAGAAATTCATTTCAAGAGCAGCAACAGCAACAGCGGTTAATCAAAAG 4138  
4262 GTTCTGAGACGGGATTAGAATAATTCAGAGGGCTGTTCTGTCAATTAAGTTGACATG 4321  
4139 CTAAAAATGATGTGTATCAAGCTGTGACAACTCAAAATCAAGCAATTTGATTAACAATG 4198  
4322 AGGTGAGCTGACGCTTGAATTAATCTTCATGGAAGAGCTAATGAATCAATGTCAATG 4381  
4199 GTGCTACAACTGGAAGAAAATGACGAAAAGATTTAGTTTAAAGCTTAAAGAAAAG 4258  
4382 CCAAGATGTCAAGAGAGTTCGTATCTAGCAAAAGATATCAGGCTTTACTAATGAA 4441  
4259 CGATCAAGATATCTTAATGACCAAACTAATGATTTACGCAAAATTTAAAGATCAAG 4318  
4442 AAGATAAAAATATTTAGAAAGATCGTGTATTAATACGACTGGAATGCTTATTATACGA 4501  
4319 CAGTCTGATATTTCAAGGTAATTAATGACAGATCAACAATTAAGATGTTG-----CGA 4372  
4502 AGGAACAATGAAAAAGCAAAAGAAAAGAGACCGGTCAATTTGTAATGCTGCTTTAT 4561  
4373 AAGATGAATTTAGCAAAAAGCAAAAGCAAAAGCGCTTATTTGCAAACTGCGAGATG 4432  
4562 CGGTCCTGGAACGATTAATCCGCTGAGAGTACTATTCAGTCAATCTACTGTTAAAA 4621  
4433 CGACTCTGAAGAAAAGAAAGCAAACTAATCAACAATGACGCAATTTAAACAAGATTA 4492  
4622 ATTAATTTAAAGCAATTTAGTGAAGCAATTAAGAAAGCGGAGAGATTAATTCATG 4681  
4493 ATCAAAATATTTGAATAATGCAACATCATCATGATGTAACACTGCAAAAAGATTAATGCA 4552  
4682 CGAAACATGTAATGTGAGAGCAAAATCATCTATCTGTGTGTGATGAGCGGCTTGGAC 4741  
4553 TTCACCAATGCAACCAATTCAGACATCAACAGATGTTAAACGAATGCAAGAGC--GGA 4611  
4742 TTGCTATCAGCAAGATGCTTTTTCAGGAATGGGATCTGAGAGCTGAGCACTTATCA 4801  
4612 TTGCTAATCTGAATGCAAAATTAATAATTAATGAAATCTTAATTAATGAGACT----- 4665  
4802 ATGACACGATTTGCAAGAGTGATTAAGAAAGAAATTTCTGCTGATTCCTTAATGTGAACG 4861  
4666 -----ACTAATGAAGAAAAGGTACGATTTGGAACAGTTAGAGCAGCATATGAAGAG 4720  
4862 CAAATTAATTCATTTCTGGGGTGAATGTTCGGGAAACATTTGCCGGTCTCTTTTCAAGG 4921  
4721 GTTTAATATATATTAATGACAACTACTACAGGTGATTAATCTGCTTAAGATACAG 4780  
4922 CGGTGAGAGCTGCTTTTTCGAAATTAATCTCTTCAATAATAAACCTCTGCTTTGATTACAG 4981  
4781 CAGTACAAAAGTTCAACAACCTTCATGCAATCTCTGTTAAGAAACGACGAGTAAAAAG 4840  
4841 AATTAGATCAAGCTGAGCTGATTAAGAAACAAATTAAGAAACAAACCAATATGCATCAC 4900  
4902 ATTCTCATTTTACAAAGTTTCTGCTGAGAGCGCTGCAAGTATTTAGCAGCTGGAATCG 5101  
4901 AACAGAAATTTAATGATGCAAAACAAAGAGTTGATTAATCAATCAAGCGAAGCAAA 4960  
5102 GAGGAATGATCTGCTGATCGTGTCTGATGAAACGGAAGCTTTAGTATGATCTG 5161  
4961 ATGTGATCAATCATCAACAATGAATATGTGATTAATGCAAGTTAAAGAAAGAAAGCTTA 5020

5162 AGTTGAAGAGTAGTACTTTCATGTAGATGCAAAAGATCAAAAACAAATAATATACAA 5221  
5021 AATTTAATGCACTTTAAACATTTAGTAGTACAAAAAGATCTTACTTAATTTGAAG 5080  
5222 TTGCGGAATGCAATGAGAGAAAAGCGCTGAGGTGGAGCAACAGTTGCTCATCA 5281  
5081 ATGCATTAATGCTTAAGTAAGAAAGAGGATTAATCTTAACCATTCGACTTCAAGTGA 5140  
5282 ATATTGAAAACATTCAGTTATAGTATTTGTAATAAAACGTAAATTTACACGGCGAATG 5341  
5141 TTGCTGAAGCGAACAACAAACCTTGCTGAATTTAAACAACTCGCATCAAAATGTTAAATC 5200  
5342 ATCAAGATTAAGAAAATATCAATGTGACTGCAAAAGATTAATTAATGACCAATTAATG 5401  
5201 AAGCTACTTCTAAGATGACATTTGAAGTTCAAAATTCATTAATGACTTAATTAATG 5260  
5402 CAGTGGAGTTGAGAGAGCAAAAGAGCGCTGTGCAAGAGCTTTCGAATTAATCACT 5461  
5261 ATTAACAATTCACACAGGTAATAAAGATCAGTACCAACAGATTTATATGCTTATGACAG 5320  
5462 TGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGAACAAGATTTAGAGAA 5521  
5321 ATCAGAGAAAATATATATATTAATGAGTGAACATAATGCAACAGATGAAAAGCAACAG 5380  
5522 AAAAAATGGAATTAAGAAAAGCAAAATGTTAATGTTCTAGCTGAAAATTAACAGTCAAG 5581  
5381 CAATTAAGCAAGTTGACCAAAATG---TTCAAACTGCAATTTAGAAAGATTAATATG 5437  
5582 TGATCAAAATGCGACAGTCTTTCGAGAGCAAGTGAACAGCTGAGTGAAGGCTGAG 5641  
5438 TGATTAATGTGATCGTGTGATGATGATTAACCAAGTAAACAGCAATTTGATGCTAATTC 5497  
5642 TAGAGTTAATAAATTAACAAATAATCTTCTGCAATTAATAAATAATTAATCTCAAAATG 5701  
5498 AAGTATGATCTACTGTTAAACCTTAAGCAACCAAGCTATTTGAAGTTAAAGCAAGATA 5557  
5702 TAGAATGCTTTGTAATAAG-----CAAAATGCTAATCACTTA 5740  
5558 CGAAAGATCTAATGATCAAAAGTGACCAAGTTAACTGCTGAAGAAAATGTAAGCATTAAG 5617  
5741 TTAAAACAATTTGGAATTTGAGCTGAGTGTGAGCTGAGAGCTGAGTGAACAGTCTTG 5800  
5618 CAATGATTTAAACAATTTACAGATCAAGTAAACAGGTAATTTACTGATGCAACAACATG 5677  
5801 TAGCAGTAATAAGATTGTAATTAATACAGATGACAGATTTAAATCTGCAAAATCACTG 5860  
5678 CTGAAGTTGAAAAGCAAGAGCTCAAGAGCTTGAAGCATTTGATTAACATTCAAATGACT 5737  
5861 CGAAGGAATGTGCAAGTTATTTACAGAGTCTGATGCGGTAATTTGTAATTAATGACGA 5920  
5738 CAACAGAAAACAAAAAGCTATCGAAGATTAAGAACTGCACTAGACCAAGATTAAGCAG 5797  
5921 CAGTGTGGAAGTGGCCGCTGAGCAATAGAGAGCTCAACAGTGTGAATGAATTAACAG 5980  
5798 GTGTAAATGTCAACGCTGATGCTACAACTGAAGAAAAGAAAGAGCTTTAAGAAATGCTTTAG 5857  
5981 GATCTACAAAAGCATATGTAAGATTTCTACAGTGAATGTGCTAAGAAAGAAACAGATATT 6040  
5858 AAGACATTTTATCAAAAAGCACTGAAGATATTTCTGATCAACTACAAATGCGAAATCG 5917  
6041 ATTTACTACTCAAGGCAAGTGAATTAAGTGTATTAAGTAATTAAGTAATTAATTAATTA 6100  
5918 CTACTGTCAAAAATTAATGTGCGTTGAACCACTTAACCAACAGTATTAATCTGAAAGTTA 5977  
6101 TTAAAGAACTATCACAAAAAGAAAATTAAGTATTAATAAAGATTTGTTTACCATTA 6160  
5978 AGAAAAATGCTTTGGAAGCAATCAAGAAAGTGTTAACAAGCAATTAAGAAATTAATTA 6037  
6161 GTTCAGCTACTCACTT 6178  
6038 ATGCAGATGCAATGAT 6055



```

QY 4982 GACGAGTAATCTTTTGTGGAAGAAATACAAAGTCAATGTACAGTTTGATG 5041
DB 4841 AATTAGATCAAGCTGCAGCTGATTAAGAAACAAATATGAAACAAACACCAATGATCAC 4900
QY 5042 ATTCTATATTACAAAGCTTTCTGCTGAGAGCGCTGCAAGTATTAAGCAGGCTGATCG 5101
DB 4901 AACAGAAATTTATATGCAAAACAGAAAGTTGATACGAAATTAATCAAGGAAACAA 4960
QY 5102 GAGGAATGATCTGTCAATCGTGTCTGATGAAACGGAAGCTTTAGTGTGTTCTG 5161
DB 4961 ATGTGATCAATCAACAAATGAAATGTTGATATGCAAGTTAAGAGGAAAGCTA 5020
QY 5162 AGTTGGAAGAGTAACTTTCTTCATATGATGCAAAAGATCAAAATCAATTAATACAA 5221
DB 5021 AAATTAATGCAAGTTAAACATTTAGTGAATCAAAAGATGCTTTAGCTAAATTTGAAG 5080
QY 5222 TTGCGGAAATGCAATGAGGAAAGCGGCTGAGAGTTGAGCAACAGTTGCTCATACAA 5281
DB 5081 ATGCAATATATGCTTAAAGTAAACGAAAGCGGATTAAGTCAAGCTGAGCTTCAAGTAAA 5140
QY 5282 ATATTGAAAAACATCAGTTATAGCTATTGTAATAAAACAGTAAATTAACACGCGCAATG 5341
DB 5141 TTGCGAAGCGAAACAAAACTTGCTGAATTAACAACTCGGATCAAAATGTTAAATC 5200
QY 5342 ATCAAGATGAAAAATATCAATGTGACTGCAAAAGATTATCTATGACCAATACTATTG 5401
DB 5201 AAGCTACTTCTTAAGATGACATTTGAAGTTCAAAATTCATATATCACTTAATTAATTAAG 5260
QY 5402 CAGTGGAGTTGAGAGGAAAGAGCGCTGTGCAAGAGCTTCTGCAAGTACTACT 5461
DB 5261 ATTACCAATTCACAGGTAAAGAAATGAAATGAGTACACAACTTTATATGCTTATGCAAG 5320
QY 5462 TGAATTAAGACAGTTTCTTCTCATGTTGATCAACCTGATTAATGACAAAGATTTAGAGAG 5521
DB 5321 ATCAGAGAAATTAATATTTTACGCTGACACTTAATGCAACAGATGAAAAAGCAACAG 5380
QY 5522 AAAATTAATGAAATTAAGAAAAAGCAATGTTATGTTTACTGAAAAATACAGAGTCAAG 5581
DB 5381 CAATTAAAGCAAGTTGCAAAATG--TTCAAACTGCAATTAAGAAAGCATTAATTAATGAGT 5437
QY 5582 TGGTCACAAAATGCAAGTCTTTCCGAGCAAGTGAACAGCTGAGTGAAGAGCTGAG 5641
DB 5437 TGGATTAATGTAACGTTGATGATGATTAACAGAGTAAAGAGCAATGATGCTATTC 5497
QY 5642 TAGCAGTTATTAATTAACAAAAATCTTCTGACATATTAATAAATAGTACTCAAAATG 5701
DB 5498 AAGTATGATCTACTGTTAAACCTTAAGCAACCAAGCTATTTGAAGTTAAAGCAAGATTA 5557
QY 5702 TAGCAAAATGCTTTGTAAAG-----CAAACTCATTCATCTTA 5740
DB 5558 CCAAAAGATCTATTGATCAAGAGCAAGTAACTGCTGAAGAAAACTGAAGCACTTGA 5617
QY 5741 TTTAAACATTTGGAATTTGAGAGCTGAGGTTGAGAGCTGAGAGAGTGAAGAGCAAGTTTCG 5800
DB 5618 CAATGTTTAAACAAATTAACAGATCAAGCTTAACAAAGTATTAATCTGTCGAAACCAACTG 5677
QY 5801 TAGCAGTAAATGATTTGTAATTAATACATAGCAGAAATTAATCAATGCAAAATCACTG 5860
DB 5678 CTGAAGTTAAAAAGGAAAGCTCAAGAGCTTGAACATTTGATTAACATTTCAATGAGACT 5737
QY 5861 CGAAGGAAATGCGAGATTATTACAGAGTCTGATCGGTAATTTGCTTAATTAATGACAGAA 5920
DB 5738 CAACAGAAAAACAAAAAGCTATCGAAATTAAGAAATCACTGACATGACAGATTTGAAGCAG 5797
QY 5921 CAGTGTCTGAGAGTCCCGTGCAGCAATGAGAGCTCAACCAATGGAATGAATTAACAG 5980
DB 5798 GTGTAAATGCAACGCTGATGCTAACATGAAAGAAAAAGAGGTTTACGAATGCTTTAG 5857
QY 5981 GATCTCAAAAAGCATATGTTAAAGATTTCTACAGTGAATTTGTTAAAGAAAGCAAGATGAT 6040
DB 5858 AAGACATTTTATCAAAAAGCACTGAAGATATTTCTGATCAAACTACAAATGCAAGAAATGCG 5917

```

```

QY 6041 ATATTACTCAAGGCAAGTAGATTAAGTGTAGATTAAGTATTCAAAAATCTTAATA 6100
DB 5918 CTACTGTCAAAAATAGTCCGCTTGACACACTTAAGCAACAGATTAATCTGAAATTA 5977
QY 6101 TTAACGAAGACTTATCACAAAAAAGAAATTAAGTATTAATAAAGAGATTTGTACCAATA 6160
DB 5978 AGAAAAATGCTTTGGAAGCAATCAGAGAAAGTGTAAACAAATTAAGAAATTAATAAA 6037
QY 6161 GTTCAAGTACTCATACTT 6178
DB 6038 ATGAGATGATGATGAT 6055

RESULT 19
ACF73459
ID ACF73459 standard; DNA; 7434 BP.
XX
XX ACF73459;
XX
XX 20-NOV-2003 (first entry)
XX
XX Staphylococcus aureus DNA #1139.
XX
XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX enzymatic assay; antibiotic target; gene; ds.
XX
XX Staphylococcus aureus.
XX
XX WO200294868-A2.
XX
XX 28-NOV-2002.
XX
XX 27-MAR-2002; 2002WO-1B002637.
XX
XX 27-MAR-2001; 2001GB-00007661.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Masignani V, Mora M, Scarselli M;
XX
XX WPI; 2003-120786/11.
XX
XX P-PSDB; ABW71899.
XX
XX New Staphylococcus aureus protein, useful as a vaccine for treating or
XX preventing Staphylococcal infection, specifically an infection caused by
XX S. aureus, e.g. sepsis.
XX
XX Claim 6; SEQ ID NO 2277; 49pp; English.
XX
XX The invention relates to novel genes and encoded proteins from
XX Staphylococcus aureus. A composition comprising the S. aureus protein, a
XX nucleic acid encoding the protein, or an antibody to the protein, is
XX useful as a pharmaceutical, particularly as a vaccine for treating or
XX preventing infection due to Staphylococcus bacteria, specifically an
XX infection caused by S. aureus. The composition is particularly useful for
XX treating or preventing sepsis in a patient. The composition can also be
XX used for diagnostics. The protein is also used in an assay for enzymatic
XX studies and as a target for antibiotics. This sequence represents one of
XX the novel S. aureus genes of the invention
XX
XX Sequence 7434 BP; 3195 A; 1326 C; 1300 G; 1613 T; 0 U; 0 Other;
XX
Query Match 0.8%; Score 79.2; DB 8; Length 7434;
Best Local Similarity 41.3%; Pred. No. 4,4e-05;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;
QY 3782 TTCAAGTAAACAGTGGCTTTGTTAAATAATGATTTTGAATGCAATTAATTA 3841
DB 3659 TTAATTAATAAGCACTGACGTAATGAATACAGCAATTTTAATTAACAAATTTGCAAG 3718
QY 3842 AAGTAATGCTTTGAGTGAAGAACTCAAGTACAGCAGACAGAGCTTTGGAAGCATTA 3901
DB 3719 AGATTCAAGTACGCGCAGATGCAACGATGAAGAAACAAACAGCTGATGCTGAAGCAA 3778

```

QY 3902 AAGAAAGTGAAGACAGAAAAAGTTATCTATTGGGAACTTCTGCTTCTATCAACTAG 3961  
DB 3779 ATACTGAAATATGTTAAAGCAATTCAGGCATTTTCAGCGACAACTCTAAACGACAAAGTTG 3838  
QY 3962 TGAACAATGAAGTTTCTGCAAAATCAGAAATTAATACAGTAGACAGAGATCTGAAGCC 4021  
DB 3839 ATGAAGCTAAAGCAATGCGAAGCAGCATTAATGCGGTAAACACAAAGTTGTGAAGA 3898  
QY 4022 AAAAAATGATGTGATGCTACTGCTTATCAAGCGGACACCCCAAGTGA CAGAGCTTTAA 4081  
DB 3899 AACAGCGGCTTAAGATGAATTTGATTCATTACAAAGCAACGCAACAAATGTTATCAATA 3958  
QY 4082 ATTTCAGAGCTGGAAGTCAAAATGGAATCTGTAGGGGCTACTGTGACTGTGGCCAATTAA 4141  
DB 3959 ATGATCAGAAAGCTCAACAGAAAGAAAAAGAGCAGCTATTCAACAATTAGCAACAGAG 4018  
QY 4112 ACAACAAGTAATGCTCTTATTAGTGGTGGAGATATATCTAACGTTAATCGAGCGAGC 4201  
DB 4019 TTACAGACGCGAAAAATATATATTACAGCTGCAACTGATGATATAGTGTAGATCAGCGCA 4078  
QY 4202 CAAAAGCTCTTTAGCAACCACTCAAGTGACTGCTGCGAGTGA CAGACGGAGGACAATTA 4261  
DB 4079 AAGACGCTGGAAGAAATTCATTAATTCAAAGCAGCAACGCAACAGCGTTAAATCAATG 4138  
QY 4262 GTTCTGAGCGGGATTAGGAATTAATCAAGGGGCTGTTCTGTCAATAGATTGCAATG 4321  
DB 4119 CTAAAAATGATGTTGATCAAGCTGTGACAACTCAAAATCAAGCAATTTGATTAATACACTG 4198  
QY 4332 ACGTGAAGCTAGCGTTGATTAATCTTCATCGAAGAGCTATGAATCAATGTCATTCG 4381  
DB 4199 GTGCTACAACTGMAAGAAAAATGCAAGCAAAAGATTAGTTTAAAGCTAAAGAAAAAG 4258  
QY 4382 CCAAAAGTGTCAAGGAAGTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGAA 4441  
DB 4259 GGTATCAAGATATCTTAAATGCA CAAACATTAAGATGTTACGAAATTTAAAGATCAAG 4318  
QY 4442 AAGATAAAAAATATTAGAAAGATCGTGTATTATATACGACTGGAATGTTATTATACGA 4501  
DB 4319 CAGTTGCTGATATTCAAGGTATTACTGCGAGATCAACAATTAAGAATGTTG-----CGA 4372  
QY 4502 AAGAACTAGTAAAAAGCAAAAGAAAAAGAGAGCGGCTCATTTGTAATGCTGCTTAT 4561  
DB 4373 AAGATGAATTAGCA CAAAAGCAAGCAAAAGAGCGCTTATTGCACAACTGCGAGATG 4432  
QY 4562 CGTTGCTGGAAGGATAATCCGCTGAGAGAGTACTATTGCACTCATATCTGTTAAA 4621  
DB 4433 CGACTACTGAAGAAAAAGAACAGCAAAATCAACAGTAGACGCAACAATTTAACACAAGTTA 4492  
QY 4622 ATAAATTTAAACAGATTTAGTGAAGCAATTAAGAAAGCCGAGAGATTAATTTCAATG 4681  
DB 4493 ATCAAAATATTGAAAAATGCAAGCTCATGATGATGTAACACTCCAAAAGTAATGCAA 4552  
QY 4682 CGAAACATGTAATGTGAGGCAAAATCATCTACTGTTGTTGTAATGCGGCTTGTGAC 4741  
DB 4553 TTCAAGCAATTAACCAATTTCAAGCATCAACAGATGTTAAACAGATGCAAGAGC-GGAA 4611  
QY 4742 TTGCTATCGCAAAAGTCTTTTTCAGGAGATGGGATCTGAGACATGCGCAAGCTTATCAA 4801  
DB 4612 TTGCTAACGGAATGCAAAATTAATACTGAATACTTAATTAATGAGACT----- 4665  
QY 4802 ATGACACGATTTGCAAAAGTGTGATTAAGAAAGAAATTTCTGCTGATTTCTTAAATGGAAG 4861  
DB 4666 -----ACTAATGAAGAAAAAGGTAAAGATATTGAGCAAGTTAGACAGATATGAAGAG 4720  
QY 4862 CAATTAATTCATTTCTTGAGGATGTTGCGGGAACCAATTCGCGGTTCTCTTTTACAGG 4921  
DB 4721 GTTAAATATATATTATGACGCAACTACTACAGGTATGTAATCTGCTTAAAGATACAG 4780  
QY 4922 CCGTAGAGCTGCTTTTGGGAATTAATCTTCTTCAATATAAACTCTGCTTTGATTTACAG 4981  
DB 4781 CAGTACAAAAGTTCAACAATCTTCATGCAAAATCTGTTAAAGAAACGACAGGTAAAAAAAG 4840

QY 4982 GAACGAAGTAAATCTTTTATAGTGAAGAAATACAAAAAGTCAATGTACAACTTTGAATG 5041  
DB 4841 AATTAGATCAAGCTGACAGCTGATTAAGAAAAACAAATATGAACAAACCAATATCATAC 4900  
QY 5042 ATTCTCATTTTACAAAGTTTCTGCTGAGGCGCTGCAAGATTTAAGCAGCTGGAAATCG 5101  
DB 4901 AACAGAATTTAAATGAATGCAAAACAAAGAAAGTTGATCTGAATTAATCAACGAAAAACAA 4960  
QY 5102 GAGGAATGATCTGTCATCTGTCATCTGATGTTCTGATGAAACGGAAGCTTTAGTATTTCTG 5161  
DB 4961 ATGTGCATCATTCACAAATGAATATGTTGATTAATGCAAGTTAAAGAGAAAAAGCTA 5020  
QY 5162 AGTTGAAGAGTAAGTTCTTTCAATGTAGATGCAAAAGATCAAAAAACATTAATACAA 5221  
DB 5021 AAATTAATGCAAGTTAAACATTTTAGTGTGACAAAAAGATGCTTAGCTTAAATTTGAAG 5080  
QY 5222 TTGCGGGAATGCAATGAGGAAAAAGCGCTGAGAGTTGAGCAACAGTTGCTCATACA 5281  
DB 5081 ATGCATATTAATGCTAAAGTAAACGAGCGGATTAACCTAACGCAATCGACTTCAAGTGAA 5140  
QY 5282 ATATTGAAAACAATCAGTTATAGCTATTTGTAATGTAATTAACAGGCGCAATG 5341  
DB 5141 TTGCTGAAGCGAAACAAAACCTTGTGATTTAAACAACTCGGATCAAAATGTTAATC 5200  
QY 5342 ATCAAGATAGAAAAAATATCAATGTGACTGCAAAAGATTATATACAACTATATAG 5401  
DB 5201 AAGCTACTTTAAAGATGACATTTGAAGTTCAAAATTCATTAATGACTTATATATTAACG 5260  
QY 5402 CAGTGGAGTTGAGAGCAAAAGAGCCTCTGTGCAAGAGGCTTTGCAAGTACTACT 5461  
DB 5261 ATTACAAATTCCAACAGGTAAAAAGAAATCAGTACCAAGATTTATATGCTTAATGAG 5320  
QY 5462 TGAATTAAGCAGTTCTTCTCATGTGTGATCAAACTGATATTTGACAAAGATTTAAGAGAA 5521  
DB 5321 ATCAGAAAGAAAAATTAATTTTCACTGACACTAATGCAACCAAGATGAAAAAGCAACAG 5380  
QY 5522 AAAAAATAGAAATTAAGAAAAAGCAATGTTAATGTTTACTGTAATAATACAGTCAAG 5581  
DB 5381 CAATTAAGCAGTTGACCAAAATG---TTCAACTGCTATTGAAGCACTTAATATAGTGTG 5437  
QY 5582 TGTGTCAAAATGCAACAGTCTTTCCGAGAGCAATGGAACAGCTGCACTAGAGACTGGAG 5641  
DB 5438 TGGATTAATGAGGAGTGTGATGATGATTAACCAAGGTAAAGCAGCAATTTATCTAATTC 5497  
QY 5642 TAGCAGTTAATTAATTAACAAATACTCTGCAATATTAATAATTTAGTACTCAAAATG 5701  
DB 5498 AAGTAGATGCTACTGTTTAACTTAAAGCAACCAAGCTATTGAAAGTTAAAGCAGAAAGTA 5557  
QY 5702 TACGAATGCTTTGGTAAAAAG-----CAATCTCATTCATCTA 5740  
DB 5558 CGAAAGATCTTATTGATCAAAAGTGACAGTTAATGCTGAAGAAAAAACTGAAGCATTTAG 5617  
QY 5741 TTTAAACAAATTTGAATTTGAGAGCTGAGAGTGTGAGAGCTGAGAGTGA CAGGTTTGTG 5800  
DB 5618 CAATGATTTAAACAAATTAACATCAAGCTAAACAAAGGATTAATCTGATGCAACCAACTG 5677  
QY 5801 TAGCAGTAATTAAGATTGTAATTAATACATAGCAGAAATTAATCATGCAAAATCACTG 5860  
DB 5678 CTGAAGTTGAAAGAACGAAACCTCAAGGACTTGAAGCATTTGATTAACATTCAATTCGACT 5737  
QY 5861 CGAAGGAATATGTCGAGATTATTA CAGAGTGTGATCGGTAAATGCTAATTTATGCAAGAA 5920  
DB 5738 CAACGAAAAACAAAAGCTATCGAAGAAATTAGAAATCGCACTGACCAAGATTTGAAGCAG 5797  
QY 5921 CAGTGTCTGAGTGGCCCGTGCAGCATAGGAGCTCAACAGTGTGAATGAATTTACAG 5980  
DB 5798 GTGTAAATGTCAAGCTGATCTCAACTGAAGAAAAAGAGCGTTTAAGATGCTTTAG 5857  
QY 5981 GATCTACAAAGCATATGTAAGATTTCTACAGTATGCTTAAAGAAAGAACAGATGATT 6040  
DB 5858 AAGACATTTTATCAAAAGCACTGAGATATTTTCTGATCAAACTACAAATGCAAGAAATCG 5917  
QY 6041 ATATTACTACTCAAGGGCAAGTAGATAAAGTGTAGATAAGTATTCAAAAATCTTATA 6100

DB 5918 CTACTGTCAAAAATGTCGGCTTGAACAACCTTAAGCCACGATTTATCTCTGAAGTTA 5977  
QY 6101 TTAACGAAGACTTATCACAAAAAATATGATATATAAAGAGATTGTTACCAATA 6160  
DB 5978 AGAAAAATGCTTTGGAACCAATCAGAGAGAGTGTATACAGCAATATAGAAATTTAAAA 6037  
QY 6161 GTTCAGCTACTCATACTT 6178  
DB 6038 ATGCAGATGCAGATGCAT 6055

RESULT 20  
AAS55232  
ID AAS55232 standard; DNA; 7437 BP.  
XX AAS5232;  
AC AAS5232;  
XX 13-FEB-2002 (first entry)  
DT 13-FEB-2002 (first entry)  
XX  
DE Staphylococcus aureus DNA for cellular proliferation protein #1544.  
XX  
XX Antisense; ds: prokaryotic cellular proliferation gene; antibiotic;  
KM antibacterial; drug design.  
XX  
OS Staphylococcus aureus.  
XX  
PN WO200170955-A2.  
XX  
XX 27-SEP-2001.  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US009180.  
XX  
PR 21-MAR-2000; 2000US-0191078P.  
PR 23-MAY-2000; 2000US-0206848P.  
PR 26-MAY-2000; 2000US-0207727P.  
PR 23-OCT-2000; 2000US-0242578P.  
PR 27-NOV-2000; 2000US-0253625P.  
PR 22-DEC-2000; 2000US-0257931P.  
PR 16-FEB-2001; 2001US-0269308P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX  
XX WPI: 2001-611495/70.  
DR P-PSDB; AAU37373.  
XX  
XX  
PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.  
XX  
XX  
PS Claim 27; SEQ ID NO 8669; 511pp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

SEQ Sequence 7437 BP; 3199 A; 1326 C; 1298 G; 1614 T; 0 U; 0 Other;  
Query Match 0.8%; Score 79.2; DB 4; Length 7437;  
Best Local Similarity 41.3%; Pred. No. 4,4e-05;  
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;  
QY 3782 TTCAAGTAAACACAGTCTTGGTTAAAAATAGATATTCGAATGCAATTAATTTA 3841  
DB 3659 TTAATTAACCACTGACAGTATGAAATCAACGCAATTTTATTAACCAATTTGCAAG 3718  
QY 3842 AAGTAAATGCTTTGAGTGAAGAACTCAAGTAGCAGCAGAGAGGTTTGAAGCAGTTA 3901  
DB 3719 AGATTCAAGCTAACGCGACAGATCCACAGATGAAGAAAAACAGCAGCTGATGTAAGCA 3778  
QY 3902 AAGAAAGTGAAGACAGAAAAAGTTATCTATTGGGAAGCTTCTCTATCAACTTAG 3961  
DB 3779 ATACTGAAAAATGTTAAAGCAATCAAGCAATTTTCAGCAGCACTACTAACGCACAGTTG 3838  
QY 3962 TGAACAAATGAAGTTCTGCAAAATCAGAAATTAATATACAGTACAGAGATCTGAAGCC 4021  
DB 3839 ATGAAGCTAAAGCAAAATCAGAGAGCAGATTAATGCGTTAACACCAAAAGTTGAGAA 3898  
QY 4022 AAAAAATGATGTTGATGTCATGCTTATCAAGCGACACCCAGTGA CAGAGCTTTAA 4081  
DB 3899 AACCAAGCGCTTAAGATGAATTAATCAATTATCAAGAACCAACCAATGTTATCAATA 3958  
QY 4082 ATTTAAGCTGGAAGTCAAAATGGAAGTGAAGGCGCTACTGTACTGTTGCCAATTAA 4141  
DB 3959 ATGATCAAGACGCTCAACAGAAAAAGAAAGCAGCTATTTCACAAATTAGACGACGAG 4018  
QY 4142 ACAACAAATGAATTCCTCTATTAGTGGGAGATATACTAGCTTAATGAGCGAGCAG 4201  
DB 4019 TTACAGACGCAAAATTAATTAATACAGCTGCACTGATGATATATGTTAGATCAGCGCA 4078  
QY 4202 CAAAAGCTCTTTTGAACAACCTCAAGTACTGCTGACAGTGAAGGAGGACAAATTA 4261  
DB 4079 AAGAGCGTGAAGAAATTCATTCAAAGCAGCAACGCAACAGCGGTTAAATCAATG 4138  
QY 4262 GTTCTGAGCGGAGTTAGAAATTAATCAAGGCGCTGTTCTGCAATAAGATTGACATG 4321  
DB 4139 CTAAAAATGATGTTGATCAAGCTGACAACTCAAAATCAAGCAATTGATTAACACATG 4198  
QY 4322 ACGTGAAGCTAGCGTTGATTAATCTTCATGAGAGAGCTAATGAATCAATGTCATG 4381  
DB 4199 GTGTTACAACTGAAGAAAAATGACAGAAAAAGATTGTTTAAAGCTTAAGAAAAAG 4258  
QY 4382 CCAAGATGTCAAAGAAAGTCTGATTCAGCAAAAGAAATATCAGGCTTTACTTAATGAA 4441  
DB 4259 CGTATCAAGATATCTTAAATGACAAACAACTAATGATGTTACGCAATTAAGATCAAG 4318  
QY 4442 AAGATTAATAATATTAGAAAGTCGTGATTAATACGATGGAATGTTATTATGCA 4501  
DB 4319 CAGTTGCTGATTAATCAAGTATTAATCTGCAATACAACTAATTAAGATGTTG-----CGA 4372  
QY 4502 AGGAACAATGAAGAAAGAAAGAAAGAAAGAGGCGGCTCATTTGAAATGCTGTTAT 4561  
DB 4373 AAGATGAATTTACCAACAAAAGCAACGAAAGGCGTTATTGACAACTGCGAGATG 4432  
QY 4562 CGTTGCTGAGACGATTAATCCGCTGAGAGTATGCTTATTCAGTCAATACGTGTTAAA 4621  
DB 4433 CCACTACTGAGAAAAAGAACAGCAAAATCAACAGTATGAGCCCAATTAACAAAGTA 4492  
QY 4622 ATTAATTTAAGCAGAAATGAGTGAAGCAATACGATGAGAGCCGAGAGATTAATTCATG 4681  
DB 4493 ATCAAAATTTGAAATGACAGATCAATGATGATTAACCTGAAAAAGTTAATGCA 4552  
QY 4682 CCAAAATCTAATATGAGAGCAAAATCACTACTGTTGTTGAGAGGCTTCTGAGC 4741  
DB 4553 TTCAAGCAATTAATCCCAATTCAGAGCTCAACAGATGTTAAACGAATGCAAGC-GGA 4611  
QY 4742 TTGCTATCAGCAAGATGCTTTTTCAGATGGAATGGAATGGAATGCAATGCAATTAATCA 4801  
DB 4612 TTGCTATCAGCAAGATGCAAAATTAATACTGAATTAATTAATGAGACT----- 4665



Db		92016	AGTAGTGGCAGTAGCACTACATGACAGTAGAGTAGAGCGACGACGACGC	92075
Oy		6396	TACCATAATTATTTAAAAAGAAATTCAGAACAAGAGTTGGAAAACTCACATGCTCATAGA	6455
Db		92076	AGCAGTAGCACTGACGACGACGACGACGACGACGACGACGACGACGACTGACAGCTAGAC	92135
Oy		6456	AGGTTTTCCGAGAAAGAACTGAAAATTACAGCAGATTTTAAGCAAAGAAATTTCCCTTTTGG	6515
Db		92136	AGCAGTAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGTAGC	92195
Oy		6516	AGTCGAGTCCGACGACGCCGGGTAGAGCCGAGTGCGAGAACCGTTTCCGTAAATCA	6575
Db		92196	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGC	92255
Oy		6576	ATTTCGACGAAGAAGCGAAGTAGATGTGGAAAGACAAAAGATTTTGTTGTAATAAAAGCTGA	6635
Db		92256	AGTAGCAGCAGTAGACGACGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGAC	92315
Oy		6636	GATTACGACAAAACGTTATAGTTCTGTGGCAATTTGGAAATGCCCGAGTGGAGTGGTCGC	6695
Db		92316	AGCAGCAGCAGCAGCAGCAGTAGACGACGAGTAGGAGTAGACGACGAGTAGCAGCAGCAGC	92375
Oy		6696	AAAAGGAGCTGGGAATTTGGACAGCAGCGAGTGAGCTTACCAAGATGAATCAACACGAGAC	6755
Db		92376	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGAGTAGACGACGACGACGACGACGACGAC	92435
Oy		6756	AAGAGTAAAAAATTTCTAAATTATGACTCCAAAACAATTGATGATTAATAGCAAAAATGA	6815
Db		92436	AGCAGCAGCAGCAGCAGCAGCAGTAGACAGTAGACGACGACGACGACGACGACGAGTAGCAGC	92495
Oy		6816	GATMAAATCAGGTACTGGAATCCGTTTCAGCCGAGCTGGAATTTCTTGACGCCGAGATATC	6875
Db		92496	AGTAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTAGAC	92555
Oy		6876	TGAGTAGGTTTCTCTCAATATATTTGCAAAATTA	6908
Db		92556	AGTAGTAGTAGTACACTATATTTCTAAATTTCTTAA	92588

XX	RESULT 22
XX	AAA70259
ID	AAA70259 standard; DNA; 4677 BP.
XX	
AC	AAA70259;
XX	
DT	07-NOV-2000 (first entry)
XX	
DE	Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:392
XX	
KW	Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW	antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX	
OS	Plasmodium falciparum.
XX	
PN	W0200025728-A2.
XX	
PD	11-MAY-2000.
XX	
PF	05-NOV-1999; 99MO-US026796.
XX	
PR	05-NOV-1998; 98US-0107131P.
XX	
PA	(HOFF/) HOFFMAN S.
PA	(CARU/) CARUCCI D.
PA	(GARD/) GARDNER M.
PA	(VENT/) VENTER J C.
XX	
PI	Hoffman S, Carnucci D, Gardner M, Venter JC;
XX	
DR	WPI; 2000-365347/31.
XX	
PT	Proteins encoded by chromosome 2 of the human malarial parasite,

PT Plasmodium falciparum, useful as antimalarial vaccines and in the diagnosis of P.falciparum infection.

PS Disclosure; Page 565-566; 577pp; English.

The present invention describes proteins and their fragments (I) encoded by chromosome 2 of the human malarial parasite, *Plasmodium falciparum*. Also described are: (1) nucleotide sequences (II) encoding (I); and (2) vaccines against *P. falciparum* infection comprising (I) or (II). (I) and (II) are useful for the development of vaccines against *P. falciparum* infection. (I) and polyclonal antisera or a monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with *P. falciparum*. Furthermore, (I) (especially when they are rRNA or secreted or membrane proteins) can aid the identification of drugs to treat or prevent *P. falciparum* infection, or they can be used to identify drug resistance in *P. falciparum*. Sequencing of the *Plasmodium* chromosome 2 and the subsequent identification of proteins encoded by it will help to expand our understanding of parasite biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AA470078 to AA470287 and AB18144 to AB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification

Sequence 4677 BP; 2106 A; 402 C; 966 G; 1203 T; 0 U; 0 Other;

Query Match	0.88;	Score 76.4;	DB 3;	Length 4677;
-------------	-------	-------------	-------	--------------

Matches 1538; Conservative 0; Mismatches 2161; Indels 45; Gaps 12.

5496 TGATATTGACAAAGATTTAGAGGAGAAATAATGGAATAAGGAAAGGCAATGTTAA 5555

Db 375 TAATAAAATTAGAGAAGCCGAGATATAAAGGAAATATCTTATTAGTAATATAGA 434

5556 TGTTCAGCTGAATAATACGAGTCAAGTGGTCACAAATGCGACAGTCTTCCGAGCAAG 5615

Db 435 AGAACCAAAGAAATATTATTGACCAATTTATTAAATAATTGACAAATTCAGAAA 494

5616 TGGACAAGCTGCAGTAGGAGCTGGAGTAGCAGTTAATAAATTACACAATACTTCTGC 5672

Db 495 ACAGAAAGTGTATCAGAAATGTACAAGTCAGTGATGAACCTTTTAAATGAATTATTA 554

5676 ACATATAAAATAGTACTCAAAATGTCGAAATGCTTGGTAAAGCAATCTCATTC 5735

Db 555 TAGGTAGATGTTAATGGAAGTAAAGAAATATTTTGGAGGAAGTCAAGTTAATGA 614

5736 ATCTATTAAACAATTTGGAATTGGAGCTGGAGCTGGAGGAGCTGGAGTGACAGG 5799

Db 615 CGATATTTTAAATAGTTTAAAGTGTTCACACAGACACACACATGTTGAAGA 674

5796 TTCTGTAGCAGTGAATAAGATTGTAATAATACGATAGCAGATTAAATCATGCAAAAT 5855

Db 675 AAAAGTTGAGAAAGTGTAGAGAAATGACGAGAAAGTGTAGAGAAATGTAGAGA 734

5856 CACTGCGAAGGAATGTCGAGTTATTACAGAGTC--TGATCGCGTAATTGCTAATTA 5912

Db 735 AAATGTAGAGAAATGACGACGAAAGTGTAGCCTCAAGTGTGAGAGAAAGTATAGCTTC 794

5913 TGCAGGAACAGTGTCTGGAGTGGCCCGTGCAACAATAGGAGCTCAACCACTGTGAATGA 5977

Db 795 AAGTGTGATGAAGTATAGATTCAAGTATTGAAGAAATGTAGCTCCAACITGTTGAAGA 854

5973 AATTACAGGATCTACAAAGCATATGTAAGATTCTACAGTGA TTGCTAAGAGAACAAC 6032

Db 855 AATCGTAGCTCCAAGTGTGAGAAATTGTAGCTCCAAGTGTGTAGAAAGTGTGGCTCC 914

6033 AGATGATTATTTACTCAAGGCAAGTAGATAAAGTGTAGATAAAGTATTCAAAA 6097

Db 915 AAGTGTGAGAAAGTGTAGAGAAATGTTGAGAAAGTGTAGCTGAAATGTTGAGA 974

QY 6093 TCTTAATATTAAGAACTTATCACAAAAAGAAAAATAGTAAATAAAAAAGATTGT 6152  
Db 975 ----AAGGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAG 1030  
QY 6153 TACCAATGTTGAGCTACTCATCTTTAAATCTTTATTTGGCAATGCGCGTGGTTCCAG 6212  
Db 1031 CTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGA--AGAAAGTGTAGCTGAAATGT 1088  
QY 6213 ACAAGCCGAGTGGCAGAGACTGTTAATATCAACAAGGTTATGAGAAACAGAGCTCT 6272  
Db 1089 TGAATAAATCGTAGCTCCACAGCTGTTGAAGAAAGTGTAGCTCCACTGTTGAAGAAATGT 1148  
QY 6273 TGTAGAAAATCTATATTAATATGCAAAACATTATTTCTGTAATCAGAGATTTACAGCA 6332  
Db 1149 AGCTCAAGTGTGAAGAAAGTGTAGCTCCAGTGTGAAGAAATGTAGTTCCAACTGT 1208  
QY 6333 TTCAATCGAGTAGTAGTTCTGTGGTGTGGTGGAAATGTAGAGTAGAGAGCTTCTTC 6392  
Db 1209 TGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCGTAGCTCCAGTGTGAAGAAATGT 1268  
QY 6393 TGATACCAATATTTAATAAAAGAAATACCAAGCAGAGTTGGAA--AACTACAATGT 6449  
Db 1289 AGCTCAAGTGTGAAGAAATGTTAGCTCCACTGTTGAAGAAAGTGTAGCTCCAACTGT 1328  
QY 6450 TGATGAGGTTTGGAGAGAAAGCTGAAATTAACAGAGATTTCTAAGCAAGATTTCTTC 6509  
Db 1329 TGAAGAAATGTAGCTCCAGTGTGAAGAAAGTGTAGCTCCAACTGTTGAAGAAATGT 1388  
QY 6510 TTTTGGAGTGGAGTGCAGCAGCCGGGTAGAGCCGAGTGGCAGAAACCGTTCCGT 6569  
Db 1389 AGTTCCAACTGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAAGTGT 1448  
QY 6570 AAATCAATTTGGCAGAAAGACCGAGTAGATGTGAAGAGCAAGATTTTGTGTAAGAA 6629  
Db 1449 TGAAGAAATCGTAGCTCCAGTGTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGT 1508  
QY 6630 AGCTGAGATTAACAGAAACGTTATAGTTCTGTTCGCAATTGCAATGCGCAGTGGAGT 6689  
Db 1509 AGCTCAAGTGTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGT 1568  
QY 6690 GCGTGCAGAAAGAGCTGCAATTTGAGCAGCAGTGGCACTTACCAAGATGCAATCAACAC 6749  
Db 1569 TGAAGAAATCGTAGCTCCAGTGTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGT 1628  
QY 6750 GAGAGCAGAGTGAAGAAATTTCTAAATTAATGACTCGAAACAAAGTTAGATG-----TAAT 6803  
Db 1629 AGCTCAAGTGTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGT 1688  
QY 6804 AGCAGAAATGATGAATAAATCAGTAGTGAATCGGTTCAAGCCGAGCTGCAATTTCTTC 6863  
Db 1689 TGAAGAAATGTAGCTCCAGTGTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGT 1748  
QY 6864 AGCCGAGTATTCGAGTGTGTTCTGTCAATTAATTTGCAATTAAGTGAAGAAACAGATAT 6923  
Db 1749 TGCCTGAAAACGTTGCACAAATTTATCAGACCAATCTTTTAAGTAAATTTATAGGTGAT 1808  
QY 6924 CGATCATAGTACTTACACTCTTCTACTGATTAATGTAATAAGCTCTTAATAAATTC 6983  
Db 1809 CGAAACTAGAGAAATTAAGACAGATTAATTAATGATTAAGAAAGTGAAGAAATGTT 1868  
QY 6984 GAATTCCTTGAACAGCCGTTGAGAGCCGAGCTTTCAGAGCAGTTACCGAGTGGTTTC 7043  
Db 1869 AGTCAACCAATTAAGAAAGTGTAGAAAGAACTACAGCTGAAGTGTACTGTTAG 1928  
QY 7044 TGTTAACCTAATAATAGTTCTGTGATAGCTGCAATTAATCTGATTTGACTTC 7103  
Db 1929 TAACTATTTAGAGGATACAGAAATTAATTAATGATTAATGATTAAGAGAAATTT 1988  
QY 7104 CGTACGAGAAATGTAATGTAACGCAAGAGAGAAATTAATTAAGCAACAGCAGC 7163  
Db 1989 AGAAGAACTCCACGAAATGTTATTAAGTGCCTTTAGAAATTAACCAAGTGAAGAGGA 2048  
QY 7164 AAATGACAGAAATCGAGAGAGCAGCAATCGAGCCAAATGTCTTGTAAATTAATTTTGAAC 7223

Db 2049 AAAAGAAAGATTAATAGATGTAATTTGAAGAGTAAAGAAAGAGGTCGTACCACTTAAT 2108  
QY 7224 AGCTGTAGAAATGAAAAAATTTCTGAAGAGAAAGAAACAGAGTTTAAAACTTTAGA 7283  
Db 2109 AGAACTGTGAAACAGCAGAGAAAGAGAGCGCAAGTCAATTAACGAAATATTTGAAAA 2168  
QY 7284 CGAAGTTAACAAAGAACAGATTAATAAAGTAAATGATGCTACGAAAAAATCTTCAATC 7343  
Db 2169 TTTAGAGAAATATGCAAGTGAAGAAATGAATGAATGAAATGTTGCAAGAAATTTAGAAA 2228  
QY 7344 AGCAGTATTTCTACAGAGTACTCTGTAAAAACGAGTGAAGAGTACTCAGAGAGA 7403  
Db 2229 CGAACTGTATTTATAGTATGATTAAGTGAAGAGAAACAGTAAATTTAGCGAGAGA 2288  
QY 7404 AGGAATTAACCCATTGGAAGACTCTGCAATTAATTTGAAAAAATGTAATTAACAAC 7463  
Db 2289 AAGTTTGAAGAAACATGAAATGATTAAGCAATTTTATAGTAAATATTTGATTAATGTA 2348  
QY 7464 AGAGAACAAAGATTAATATCACTTCTACTGTGTGGTGGAACTGCAGGCTCTGCCGC 7523  
Db 2349 AGGAATACAGAAATTTT--ATTACAGGTATGTTTGAAGTATAGAAACAGATATGT 2405  
QY 7524 ATCAGAAACAGTGCAGTTACAAATATTAAGAAATTCGGAAGTACTGTTGAAAAATC 7583  
Db 2406 AATCCAATCAGAGAAAGGTTGATTTGAATGAAATGTTGATAGTTGATTTAGATTA 2465  
QY 7584 TTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGTCCGATATTAACAGAAATGTTGCTTT 7643  
Db 2466 TATGAAAAATGAAAGAGGTTTATTAATTAATTAAGAAATTTTCAAGTACTG---- 2521  
QY 7644 AACGACATCAAGTCTGTAGAGACATTTGGAAATAGAGCTGCTATGCAATTAATA 7703  
Db 2522 --AAGTGTCAAGAACTGTACTGAACATGTGAACAAATGATATGTGATGTTGA 2579  
QY 7704 TTTTAATGAAGATCAATATCAATATTAATAATTTCTAAGCTATTAAGAAAAATTTGA 7763  
Db 2580 TGTTCCTGCTATGAAGATCAATTTTATGAAATTAATAAGAGCAGAGGTTGAAGA 2639  
QY 7764 TGTATTGTAAAGATTAATGGAATTTGAGCCGAGCAAAAGATTAACCTATAGAGC 7823  
Db 2640 AATGTTTTTAATTTGGAAGATGATTTTAAAGTGAAGTATGATTAATCTGTAGAGA 2699  
QY 7824 GGTAGTCCGAGCCATTAATCTCAAAAGCAAGAAATGAA--TGAATTCAGAGTTGA 7880  
Db 2700 AATTAAAGATGAACCGGTTCAAAAGAGTGAAGAAAGAACTGTAGTATTAATGAAGA 2759  
QY 7881 AATTGAGAGATATTTCAATGAAGAAATAGATTAATAGCCCTTCTAAAGAAATGG 7940  
Db 2760 AATGAGAAATTAATTTGATGATTAATGAAGAAAGAAAGATTTTAACAGACAAAGAT 2819  
QY 7941 AAGGAAATCAATGTCAAAGTGAAGAAAGAAACAGAGTACTGTAATCTCAAGAGC 8000  
Db 2820 GATGATGCACTAGTAAGAAATCATAGAAATCTTCAGATTTCTAAAGAAACCTGAATC 2879  
QY 8001 TTTCTGAGAGCAGTACAGAGGAGGAGGAATTAATTTCCGAACCAAAAGATGCGGAAGCTC 8060  
Db 2880 TATTAAGATTAAGAAAGATGTTTCACTAGTTGTTGAAGAGTTCAAGCAATGATAT 2939  
QY 8061 TTAATTTGAAGTATGTAACAAATCCGGAAGAGATTTTTCATGCAGATTAATGTAATAT 8120  
Db 2940 GATGAAAGGTTGAGAAAGTTTATGAATTAAGAAATTTGAAAGAGGATTAAGAA-- 2996  
QY 8121 GGAAGCAACATTAATTAAGAAAGTGAACAGAGTTTCTTAAGCAGTAAACAGGTTCTGTAT 8180  
Db 2997 GGAATGCTGTGAAATTAATGATCAATTAAGCAATTAATTAAGAAATCTCAAGAGTTTAA 3056  
QY 8181 GGAAGAGTGTGAGTCAACCAAGCAGAAAGTAACTGTCGAGATTAACATATGATGAGAT 8240  
Db 3057 TGAAGTGAAGCAATTTAATTAAGATTAAGAAATTTAAAGAAATTTAGAGAAAGCAAT 3116  
QY 8241 TGAAGAGAAATTTGTTCAAGAACAAATCGATTAATGCAATTTCTAAAGTGAAGAGTTT 8300

```

Db      3117 ATCAGAAAGTTCTTAAGAAATTAATAGATGCACAAAGATGATACATTAGAAAAGTTATGCA 3176
Qy      8301 GGATGAAAGATAAGTAAGTCTGCTAAATCTTCTGATGATCAGAAAATGAGAGAAATTGC 8360
Db      3177 AGAGGAACATGATATTAACGACGAGCGTGTGATGAAAGTTAGAAATTAAGAAATGTCGAGAA 3236
Qy      8361 CGAGACGAGAGTGAATCTTCTTACAGCACAAAGTAATCTGAATCCGTAAGTTCCGTTTACG 8420
Db      3237 AGACAAAGATTCAGAAAAGTATGATTTAAAGATCTTAAGAAAGTAATTAATTAAGAAAGT 3296
Qy      8421 AAAGCAAGATTAATGAAATTAATGATTTACACAAAATAATTAATTCAGAAAGTCAATGCTCT 8480
Db      3297 AAAAGAAATCAAGAACTTGAAGTGAATTTTAAGAAATTAAGAAATTAAGAAATTAAT 3356
Qy      8481 TCCTTTAATGATACAAAGAAATGAAGCAATTAAGATCTTTAGCGGTAGCCGATGTCGA 8540
Db      3357 TGAACACAGATTTTAAAGAAAGAAAAGAAATTAAGAAAGTCAATTTGAAAATTTGCA 3416
Qy      8541 TGCACAGAAACAAACAAAGCTTTACAGATCAAAAGTTAACTTCTTACAACTGTAA 8600
Db      3417 AGAAGACGTGAAGAAATTAAGATCTTGAAGCAGATATATTAAGAAAGTAATCTTCAAT 3476
Qy      8601 TGGAGAAACGTATCTCAACTTCGTGCAAAAGCTTTGGCTAAATAAGAAATTTATGAAA 8660
Db      3477 AGAAGTTGAGAGAAAGAAAATTTAGAAAGATCACCAATTTAAAGAAAGGTAGAAC 3536
Qy      8661 TGTAAAGAACTGAGAGAGCCTTAGTCGAGCGAGAAACG-----CAGCCGTTGAAA 8714
Db      3537 TATATTAAGTGTGATGCGCATATTAAGAGTTTGAAGAAAGTGAATTAAGAAAGATGAG 3596
Qy      8715 TTATACAAAGATCTACAGAGCATGTTGTCAGAGAAATTTGGAAATTTGAGAAAT--AA 8772
Db      3597 TGAATTTAAAGAGAGATATTTAGACATGTTAAAGGAGATTAAGAAATTAAGGGAATATGCA 3656
Qy      8773 TTAGAAAGATTCGAAAGATTAATACGATTTGAAGTCAACGAGAGCAAGCAAAAGCA 8832
Db      3657 TAAAGAAATTTAGAAAGATGTAACGCAAACTTGAGAAAGAGTTGAATCTTTAAAGAA 3716
Qy      8833 GGTCTTGTGCAAGAAAGATGTAATTTCTGTAAGAAATTAATCAATTTCAAGGAGAA 8892
Db      3717 TCTTTATCTAGTGCATTTAGCGATGATGAAGAACAAATTAAGAAAGAAAGAAAGCTCA 3776
Qy      8893 TCCATTGAAGATTAAGCCAGAAATGTTGACCCGAAAGTGAATATGATGATGCTTTGAAT 8952
Db      3777 AAGACCTTAATTTGAAGAGATATTTAATAAGAGAGTTAAAGAAAGCAAAAGAAAA 3836
Qy      8953 GAACTTGAATGATCTACAGAGAAAGAGTGTGCTATGCTGGAAT----TGTATTGG 3908
Db      3837 AATTAACAAAAAGAAAGTAAGGTTGATATTAAGATTAAGAAAGCAAAAGATGAATAGT 3896
Qy      9009 AAATGTTGATTAATTAATATGATTAAGAAAAATGTAAGAACCAAAATTCGAAAGCATGC 9068
Db      3897 AGAAGTTGAAATCAAGAGATGAAGATTAATGATGAAGATTAAGAAAGAGATGTAAGAAAG 3956
Qy      9069 TATTTAGAAACTACTGAGAAACAGAAATATCAAGATTTTACAGAGCAAAAGTAAATAT 9128
Db      3957 TATAGAAAGATTAAGGTTGAAGATATAGTGAAGATTAAGATGAAGATTAAGATGAAG 4016
Qy      9129 TCTTGAAGAAAGAGAGCTGCGAGCTGACGTGAATATGCAATGTCACATTTCCATG 9188
Db      4017 TATAGTGAAGCAAAAGATGAAGTATAGATTTTAATATGTCCAAAAGAGAAAGCATTTGA 4076
Qy      9189 GATGATATTTAAATTTGGCAAA 9212
Db      4077 AAAGCTTTAAAGAGAAAAAGAAAA 4100

```

RESULT 23  
 ACL64817  
 ID ACL64817 standard; DNA; 1039 BP.  
 XX  
 AC  
 AC ACL64817;

```

DT      02-JUN-2005 (first entry)
XX      M. xanthus DNA fragment, seq id 1280.
DE      Transgenic plant; DNA replication; gene regulation; gene expression; db.
XX      Myxococcus xanthus.
XX      US6833447-B1.
XX      21-DEC-2004.
XX      10-JUL-2001; 2001US-00902540.
XX      10-JUL-2000; 2000US-0217883P.
XX      (MONS ) MONSANTO TECHNOLOGY LLC.
XX      Goldman BS, Hinkle GJ, Slater SC, Wiegand RC,
XX      WPI; 2005-028716/03.
XX      New substantially purified Myxococcus xanthus nucleic acid molecule
XX      encoding a nitrite reductase, useful for determining gene expression,
XX      identifying mutations in a gene of interest, and for constructing
XX      mutations in a gene of interest.
XX      Example 1; SEQ ID NO 1280; 25pp; English.
XX      The invention relates to a substantially purified nucleic acid molecule
XX      encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
XX      recombinant DNA construct for expression of a nitrite reductase gene in a
XX      plant cell, and a plant cell comprising the recombinant DNA construct.
XX      The nucleic acid is useful for determining gene expression, identifying
XX      mutations in a gene of interest, and for constructing mutations in a gene
XX      of interest. Sequences given in records for SEQ IDs 1-1849 represent a
XX      set of 1849 contigs and singleton sequences comprising coding sequences,
XX      DNA replication elements, promoters and other regulatory elements from
XX      the genome of the bacterium Myxococcus xanthus. Note: The sequence data
XX      for this patent did not form part of the printed specification, but was
XX      obtained in electronic format directly from USPRO
XX
SQ      Sequence 1039 BP; 819 A; 72 C; 17 G; 129 T; 0 U; 2 Other:
Query Match      0.8%; Score 76; DB 14; Length 1039;
Best Local Similarity 44.1%; Pred. No. 0.0001;
Matches 403; Conservative 0; Mismatches 507; Indels 4; Gaps 2;
Qy      144 AAAAAATTAAGAAAAAGATTAATGTTATGACATTTACTACAAACAGATTCAAGGAGAA 203
Db      107 AAAAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 166
Qy      204 CGCTTTTAAACAGTTTAAATAGATTGCTTTAAACAGAAATTAATTAAGCAATCTATATTT 263
Db      167 TTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 226
Qy      264 TGGGAGAAAGAAATAGTACGGGGGTAAATATATCTTTTAACTTTGTCAATGAAAAATTGA 323
Db      227 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 286
Qy      324 ACTAGATGGGATTTCAACGGAATTCGAGAAATTAATTAATTAATTAATTAATTTATTTCTT 383
Db      287 AAAAAAATCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 346
Qy      384 AAGCTCGAAGGATGACAGTGAAGAAATTAAGAGTATCAATGCTGCTTTTCATTC 443
Db      347 AAAAAATTAATTAATTAAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTA 406
Qy      444 TATTAATCCAAACAAAGATGATTTTAAGAAAGCTTTGAAAGAACCAACATGTTAAAGT 503
Db      407 AATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 466
Qy      504 TTTTAATGAATCATTCAGTAGAGAAAGTAATTAATTAATTAATTAATTAATTAATTA 563

```

```

Db      467 CCAAAAAAAAAAATAATAAAAAAAAAAAAAAAAAAAGAAAAACAAAAAATAAAAAAAAAA
Qy      564 CATTACGGTAGAAGAAAAATCAATGCTGTGAAGGACGCGTGTATATGCGCGGATAT 623
Db      527 AATTAATAAAAAAAAAAATAATAAAAAAAAAAATAATAA-AGATATAAAAAAAAAAAGT 585
Qy      624 TAGATTGAAGATAGTCAATCTAATAAGACAGAAATTAAGATTTTAAATAATTTAGTCA 683
Db      586 AAAATACATTAATTTAAAAAACAATAAAAAAAAAACNCGGACATTAATATAATAATAATA 645
Qy      664 TATTGTGATGCAATTAATTTCTGCTGACCGGAGATTAAAAAGCTACCAAGACAAATC 743
Db      646 CTATATAATAAACAATAAAGTGTCTTTCTTAAAAAATAAATCTCCCTAAAAACA 705
Qy      744 TGGAGATATTATTCTTTCAGCTCAGATGATTTCTCTCAAAAAGC---TATGGGAAAA 800
Db      706 TAAAAAATAAAAAAAAAAATAAATAAAGGCTCTAATCAACCCCAATATTAACAAAA 765
Qy      801 TTCAACTGTGGAAGAGATAGAAATATGTATAAAGAAATACCAAGCAATATTTGA 860
Db      766 TCTTAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 825
Qy      861 ATCTGATGCTGTATTGGAAGCAGATGGAATATTAATAATTTAGTGGAAAGCTACAAATG 920
Db      826 AACTATAAATAAATAAAGGGAATAAATAAATAAATAAATAAATAAATAAATAAATAA 885
Qy      921 GAGATTATTAAGAAGAAAGGGAATAAATAAATAAATAAATAAATAAATAAATAAATAA 980
Db      886 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 945
Qy      981 TGTGGAAGCTTCCTAGAGATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1040
Db      946 TAAAAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1005
Qy      1041 AGCTGAAGCAAGA 1054
Db      1006 AAAAAAAAAAAAAA 1019

RESULT 24
AAS51998
ID      AAS51998 standard; DNA; 7104 BP.
XX
XX
AC      AAS51998;
XX
DT      13-FEB-2002 (first entry)
XX
DE      Staphylococcus aureus DNA for cellular proliferation protein #415.
XX
KM      Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
XX      antibacterial; drug design.
XX
OS      Staphylococcus aureus.
XX
PN      WO200170955-A2.
XX
PD      27-SEP-2001.
XX
PF      21-MAR-2001; 2001WO-US009180.
XX
PR      21-MAR-2000; 2000US-0191078P.
PR      23-MAY-2000; 2000US-0206848P.
PR      26-MAY-2000; 2000US-0207727P.
PR      23-OCT-2000; 2000US-0242578P.
PR      27-NOV-2000; 2000US-0253625P.
PR      22-DEC-2000; 2000US-0257931P.
PR      16-FEB-2001; 2001US-0269308P.
XX
XX
PA      (ELIT-) ELITRA PHARM INC.
XX
PI      Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
PI      Yamamoto RT, Xu HH;

```

```

XX      WI; 2001-611495/70.
DR      P-PSDB; AAU34139.
XX
PT      New polynucleotides for the identification and development of
PT      antibiotics; comprise sequences of antisense nucleic acids.
XX
PS      Claim 27; SEQ ID NO 4580; 511bp; English.
XX
CC      The invention relates to antisense inhibitors of genes essential to
CC      prokaryotic cellular proliferation, their use in identifying the genes,
CC      their use in the discovery of novel antibiotics, the essential genes
CC      themselves and the encoded proteins. The prokaryotes used are Escherichia
CC      coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC      Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC      useful for the identification of potential new targets for antibiotic
CC      development. The antisense nucleic acids can also be used to identify
CC      proteins used in proliferation, to express these proteins, and to obtain
CC      antibodies capable of binding to the expressed proteins. The proteins can
CC      be used to screen compounds in rational drug discovery programmes. The
CC      antisense nucleic acid sequence is also useful to screen for homologous
CC      nucleic acids which are required for cell proliferation in a wide variety
CC      of organisms. The present sequence encodes an essential prokaryotic
CC      cellular proliferation protein. Note: The sequence data for this patent
CC      did not form part of the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 7104 BP; 3045 A; 1258 C; 1249 G; 1552 T; 0 U; 0 Other;
XX
Query Match      0.8%; Score 74.8; DB 4; Length 7104;
Beet Local Similarity 41.6%; Pred. No. 0.00031;
Matches 992; Conservative 0; Mismatches 1367; Indels 27; Gaps 7;
Qy      7208 TAAATATTGGAACAGCTGTAGAAAGTAAAAATTTGGAAGAAAAAGCAAGAG 7267
Db      2702 TCAATGCAATTCAACACAAACAAAGTTAAACCTGACGCTGAATCGGAATGAAAAACG 2761
Qy      7268 TTTTAAAAAAGTTTGAAGCAAGTTAAACAAAGCAAGATTAATAATGATGCTACGA 7327
Db      2762 CATATAATACGCTTAACCAAGAAATTCAAATATAGCAATGCTTCACTACAGAAAGAAAC 2821
Qy      7328 AAAAAATCTTACATACAGACAGATTTCTACAGAAAGATCTTGTAAAAAGCGATAGAG 7387
Db      2822 AAGCTGATATACAGAAATTAATATCTAAAAAGCAAGAAAGCAAAATCTTATGCTG 2881
Qy      7388 GAGATCTCAGGAGGAAGAAATTAAGCAATTTGGAAGACTTCTGATTTATTTGAAAAA 7447
Db      2882 CAAATACAAACAGTGAATGTAACAACAGCTAAAGCAATGTATGCTGCAATTAATCAAG 2941
Qy      7448 ATGTAGATATTACAAACAGAGACAAGAAATATATCACTTCTACTGTGTGTTGGAACTG 7507
Db      2942 TACAAAGCGCAACAATTAAGAAATCGGATGCTTAAGCGGAAATGGCTCAAAAAACCAAG 3001
Qy      7508 CAGGCTTGTCTCCGATCAGAAACAGTGCAGTTACAAATATTAAAGAAATTCGGAG 7567
Db      3002 AAGCTAAAACT---GCAATTGAACCAATGAATGATTCGACTACGAGAAACAACAGCTG 3058
Qy      7568 TTACTGTGAAAAATTTCTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGATCGATATTA 7627
Db      3059 CAAAAGATAAAGTTGATCAACAGTATTAATCTGCAACGCTGATATAGATATCTGCGAG 3118
Qy      7628 CAGGAATGTTGCTTTAAACAGCATATCAAGGCTCTGTAGAGCAATGGGAAATGAGGCTG 7687
Db      3119 CAAATACGATGTGATATATCAAAAACTTAATTAAGAGCTACATCGACCATTTACAC 3178
Qy      7688 CCTATGCAATTAATTTCTAATGGAAGATCAAAATATCAGTATTAAAAATTTCTAAGCTAT 7747
Db      3179 CTGATGCAAAATGTTAAACCAACAGGAAACAAGCAATGCTGATTAAGTATACAAAGCGAAG 3238
Qy      7748 TAGGAAAAAATATTGATGTTA---TTGTAAAAAGTAAATCGAANTTGAGACCGAAGCA 7804
Db      3239 AAACAGCAATTGATGCTAATTAACGATGCAACAGAAAGAAAGAAAGCAAGCTCGAAACAA 3298

```



PR 23-OCT-2000; 2000US-0242578P.  
 PR 27-NOV-2000; 2000US-0253625P.  
 PR 22-DEC-2000; 2000US-0257931P.  
 PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JM, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX

DR MPI: 2001-611495/70.  
 XX P-PSDB; AAU36795.

PT New polynucleotides for the identification and development of  
 XX antibiotics, comprise sequences of antisense nucleic acids.

PS Claim 27; SEQ ID NO 8291; S11pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC their use in the discovery of novel antibiotics, the essential genes  
 CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*  
 CC *coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
 CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence encodes an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pcc\_sequences

XX Sequence 7107 BP; 3047 A; 1258 C; 1249 G; 1553 T; 0 U; 0 Other;

Query Match 0.8%; Score 74.8; DB 4; Length 7107;  
 Blast Local Similarity 41.6%; Pred. No. 0.00031;  
 Matches 992; Conservative 0; Mismatches 1367; Indels 27; Gaps 7;

QY	7208	TAAATATTTTGGAA	CAGCTGTGAGATGAGAAAAATTTCTGAGAGAAAAAGGACAGAG	7267
DB	2702	TCATATGCAATTC	CAACAAAGAAAGTTAAACCTGCAGCTGATACGGAAATGAGAAAACG	2761
QY	7268	TTTTAAAACTTTAG	CGAAGTTAACAAAGAACAGATAAAAAGTAAATGATGCTACGA	7327
DB	2762	CATATATATACG	TAAACAAAGAAATTCAAAAATGCAATGCTTCAACTACAGAAAGAAAAC	2821
QY	7328	AAAAAATCTTAA	TCACAGCAGATTTTCTACAGAAATATCTTCTGTAAGGGGATAGAG	7387
DB	2822	AGCTGCATATAT	CAGATTTAGATTAATAAGCAAGAACAAATCTTGTATGCTG	2881
QY	7388	GAGATATCTCAG	GAGAAAGAAATTTAAAGCATTTGTAAGACTCTGTATATTTATGGAAAA	7447
DB	2882	CAAAATACAAAG	AGTATGTAACAACGCTAAAGACATGATTTCTGCAATTAATCAAG	2941
QY	7448	ATGTAGATTTTCA	ACAGAGAACAGATTAATATCACTTCTACTGTGTGTTGGAACTG	7507
DB	2942	TACAAAGCGGCA	CAACTAAGAAATCGATGCTAAAGCGGAAATCCCTCAAAAAGCAAGTG	3001
QY	7508	CAGGCTTGCTT	CCGATCAGAGACAGTGGCAGTTTACAAATATTTAAAAAAATTTCCGGAG	7567
DB	3002	ACGTTAAACCT	---GCAATTTGAAGCAGATTAATGATTTGCACTAGTAAGAACAAACAGCTG	3058
QY	7568	TTACTGTGAAAT	CTTTTGTGAAAGCAGCTGAAAAAGTAATGTTAGATCGATATTA	7627
DB	3059	CAAAAGATTAAG	TTGATTCAGACAGTAGTACTGCAAAACCGTATATAGTATATGCTGCAG	3118
QY	7628	CAGGAATGTTG	CTTTACAGCATTCAAAGGTCCTGTAGAGACATTTGGAAATAGAGCTG	7687

DB	3119	CAAACTAGATGTAG	ATATGCAAAAACCTACTAATGAAAGCTACAAATCGACCATTTACAC	3178
QY	7688	CCATATGCAGAAT	TTAATTTCTAATGGAAGATCAAAATATCAGTATTAAAAATTTCTAAGCTAT	7747
DB	3179	CTGATGCAATG	TTAAACCAACAGCGAAACAAAGCAATTTGCTGATTAAGTATACAGGCGAAG	3238
QY	7748	TAGAAAAAATAT	TGATGTTA---TTGTAAAAAGATTAATTCGAAATTTGAGACCGAAGCA	7804
DB	3239	AAACAGCAATG	ATGCTAATTAACGATGCAACAAAGAAAGAAAGACAGCTCGAAACAAAC	3298
QY	7805	AAGGATTAAC	CGTATGAGACGGTAGCTCCGAGACCATTTATCTCAAAAGCAAAAGATGAAA	7864
DB	3299	AAAGTTCAAC	CTGAAAAACAAACAGCTGATACAGCAATTTGATGTGCATCAATCAATGCG	3358
QY	7865	TGAATTCAGAG	TTGAAATTTGAGAAAGATTTTCAATGAAAGAAATAGAGTAACTAGCC	7924
DB	3359	AAAGTTAAAC	CGGCTTAATAATTCAGAAATTTGCTTAATTTGAAAGCAATTTAGCCAGCAACA	3418
QY	7925	CTTCTAAAGGA	ATTG---GAAAGAAATCAATGTCAAAGTGGAAAAAGAAAAACAGAGTGA	7981
DB	3419	CAACTAAAGAT	TAATGCGAAACAGCAATTTGCTAGGAAAGCAATGAACTGTAATTAACAGCA	3478
QY	7982	CTGCTGAAT	CTCAAGAGCTTCTGTAGACAGTACAGGCGCAGAAATTTTCCGAAAG	8041
DB	3479	TGCTTCAAC	CCCAAGACATTAATCTGTGTAAGAAATTTGCGCGCAATTCGCAAT---GTAG	3535
QY	8042	CAAAAGATCG	CGGAAGCTTATTTTGAAGTATGACAAATTCGGAAGAAAGTATTTTC	8101
DB	3536	ATAATGCTGT	ACACAGCAAAATTAACCAATTTGAAGCTCTAATATGTAATATGATGTAG	3595
QY	8102	ATGCAGATAT	GTGATATGAGAACACACATTAATAAGTAAACAGACAGTTTCTAAAG	8161
DB	3596	ACCAAGCGAAA	CAACAGCTGTAGAGTATGATTCAGAGTAAACCAACAGTATTAATAA	3655
QY	8162	CAGTACAGG	TTCTGTATTTGGAGAGATTTGGAGTCCAAAGCAGAAAGCTACTGTCAG	8221
DB	3656	AAGCAACG	CACTTACACACCGAAATAATATTTACAGCTGCAACTGATATATATGTGTG	3715
QY	8222	GTAATAATAT	GTGAAGTTGAGGAAGAAATTTGTTGAGAACAATAGATTGAATGCA	8281
DB	3716	TGATATACG	CAAGATGACAGTAAATTTGATTCAAAGTACACACACAGCAGCAGCG	3775
QY	8282	TTTCTAAAG	TGAAGTTTGGATGAGATTAAGTAACTGCTAAATCTTCTGTAGTATCAG	8341
DB	3776	TTAATCAAA	TGATTAATAATATGATGTATCAAGCTGTGACAATCAAAATCAAGCAATTTG	3835
QY	8342	GAAATGAG	GAAGAAATTTGCCGAGCAGAGTGAATCTTCTACACACAAAGTAACTG	8401
DB	3836	ATAATACAA	CTGTGCTACAACTGAGTAAAGAAAAATGCGAAAAAGATTTAGTTTAAAG	3895
QY	8402	AATCGTAG	TTGTTTGAAGCAAGTATTTGAATAATTAATGATTAACAATAAAATATA	8461
DB	3896	CTAAAGAAA	ACGATTAAGATTAATTTAAATGCAAAACAACTACAGATGTTACGCAA	3955
QY	8462	TTTCAGAA	GTCAATGCTCTCTTTAAATGATACAAAGAAATGAAGCGAATATAGATCTT	8521
DB	3956	TTAAAGAT	CAAGAGTTGCTGATGTCAAGGATTAATCTGAGATACAACTAAAGATG	4015
QY	8522	TAGCGGTG	CCGGTGTGATGACAGAGAACAAACAAAGCACTTTACGAGATCAACAACT	8581
DB	4016	TTGCGAAA	GAATTTAGCAACAAAGCAAGGAAACAAAGCACTTATTGACAAACTG	4075
QY	8582	TAACTTCA	CAACGTAAATGAGAAACGATATCACTTCGTGCAAAAGCTTTGGCTA	8641
DB	4076	CAGATGCA	CTACTGAAAGAAAGCAAGCAAAATCAAC-----AACTAATGCAAGAT	4129
QY	8642	AAAAATGA	AAATTTATGAAATGTAAGAAAGCTGAGAGACCTTAGTCGAGCGGAAACAG	8701
DB	4130	TAAAGCAAG	ATATCAAAATTTGAAATATGCAAGTCAATGATGATGTAACACTGCA	4189
QY	8702	CAGCGGTG	AAATTTATACAAAGAGTACTACAGAGCATTTGGTTGACAGAAATTTGGGAA	8761
DB	4190	AAAGTAA	TGCAATTTCAAGCAATTTGACCAATTTCAAGCATCAACAGATGTTAAAAAGATG	4249

```

QY 8762 TTGGATATAATAGAAAGATTGCAAGAGATTAATACATTGTAGAGTCAACGAGACG 8821
DB 4250 CAAGAGCGGATTTGTAATGTAATGCAAAATTAATTAATCTTAATTAATTAATG 4309
QY 8822 GAACCAAGAGAGGTCTTGTGCGAAAGAAATGTATTTCTGTGAAATAATCAATTTTCAGGGG 8881
DB 4310 AGACTACATTAATGAAGAAAGGTTAAAGATTTGGACGAGTTAGAGCTGATTAAGAGAG 4369
QY 8882 AAACAAATTCATCATTTGAAAGATTAAGCCAGAAATTTGTGAAACCGAAGTGTAAATGTAG 8941
DB 4370 GTTTAAATTAATTAATTAATGAGCAACTAGAGGTGATGTAATCTAGTAAAGATAGAG 4429
QY 8942 ATGCTTTGATTAATCTGATTAATGATCTACAGAGAAAGGTGTGCTATGTGTGAATTTG 9001
DB 4430 CAGTCAAAAGATTCACAACTTCATGCAAACTCTGTTTAAGAGCCAGACAGTAAACCTG 4489
QY 9002 GTATTTGAAATTTGATTAATTAATTAATGATTAAGAAATTAAGAAATGTAAGCCAAATCGGA 9061
DB 4490 CTTTGAATCAAGCTGAGCTGATTAAGAAACCAATTAACAAACCAATTAATGCTGAC 4549
QY 9062 GACATGCTATTTAGAACTAGTGAAGAAACAAAGAAATTCAGATTTACAGAGCAAAAG 9121
DB 4550 AACAAGAAATTAACGATTTGCAAAACAAAGAAAGTTGATCTGATTAATTAATCAAGGAAAA 4606
QY 9122 TAAATATTTCTGAAAGAGAGAGCTGACAGCTGACAGCTGCAATTCGAATGTACACATTT 9181
DB 4607 CAAATATTTGACCAATCTTCACAGAGATTAATGTTGATTAATGATTAAGAGAGGAAAG 4666
QY 9182 CCAATGAGATGATTAATTAATTAATTTGCAAGAGATGATCATCTTCAATTAATTAACCA 9241
DB 4667 CTAAATTAATGATTAATTAATTAATTAATGATTAAGAGTAAAGAAAGTCTTGAAGATTG 4726
QY 9242 AAAATTTCAAAATTAATTAATTAATTTAGCTTTAGCATC-----AAGTGAATGAAATGTGAATG 9295
DB 4727 AAGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 4786
QY 9296 TTCATGAGGTGCTGAGAGAGAGAGTGCAGAGAGCCAAAGCAGATTAATGTTAAAGATTC 9355
DB 4787 AATTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4846
QY 9356 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9415
DB 4847 ATCAAGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4906
QY 9416 ATGTAATGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9475
DB 4907 AGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4966
QY 9476 CGGATGCGCAAAAGTCAATGCTGAGCTGCTCGGCAATGCACTATTTGAAAAAATGAAG 9535
DB 4967 CAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5026
QY 9536 TAAATTTTAATTAATGAGATCCGAAATTTAAATTAATTAATTAATTAATTAATTAAT 9581
DB 5027 AAGCAATTAAGCAAGTTGAACAAATGTTCAAACTCATTAAGAGAG 5072

```

```

RESULT 26
AAV74374
ID AAV74374 standard; DNA; 8155 BP.
AC
AC AAV74374;
XX
XX
DT 16-MAR-1999 (first entry)
DE Staphylococcus aureus contig SEQ ID #63.
XX
XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KM cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KM skin infection; surgical wound infection; scalded skin syndrome;
XX toxic shock syndrome; ds.
XX

```

```

OS Staphylococcus aureus.
XX
XX Key Location/Qualifiers
XX misc_feature 1321..1380
XX /tag= a
XX /note= "these bases represent a line of missing text in
XX the sequence listing in the specification. They are
XX included to maintain the nucleotide numbering given in
XX the specification for this DNA sequence"
XX 3121..3180
XX /tag= b
XX /note= "these bases represent a line of missing text in
XX the sequence listing in the specification. They are
XX included to maintain the nucleotide numbering given in
XX the specification for this DNA sequence"
XX 4921..4980
XX /tag= c
XX /note= "these bases represent a line of missing text in
XX the sequence listing in the specification. They are
XX included to maintain the nucleotide numbering given in
XX the specification for this DNA sequence"
XX 6721..6780
XX /tag= d
XX /note= "these bases represent a line of missing text in
XX the sequence listing in the specification. They are
XX included to maintain the nucleotide numbering given in
XX the specification for this DNA sequence"
XX
XX EP786519-A2.
XX
XX 30-JUL-1997.
XX
XX 07-JAN-1997; 97BP-00100117.
XX
XX 05-JAN-1996; 96US-0009861P.
XX
XX (HMDA-) HUMAN GENOME SCI INC.
XX
XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX WPI; 1997-374922/35.
XX
XX polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX PT scored on computer readable medium and used in the production of anti-
XX S.aureus vaccines.
XX
XX Claim 1: Page 482-487; 3271pp; English.
XX
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
XX of the invention. The DNA sequences are recorded on a computer readable
XX medium, preferably selected from a floppy or hard disk, random access
XX memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
XX the S.aureus DNA sequences allows putative functions to be assigned so
XX that protein-encoding or regulatory regions of commercial, therapeutic or
XX industrial importance can be obtained. Specifically, sequences which are
XX likely to encode antigens have been identified and these polypeptides can
XX be used in a vaccine composition against S.aureus infection. The
XX CC polypeptides can also be used in a kit for the immunodetection of
XX S.aureus in a sample. S.aureus is implicated in numerous human diseases,
XX including cellulitis, eyelid infections, food poisoning, osteomyelitis,
XX skin and surgical wound infections, scalded skin syndrome, toxic shock
XX syndrome, etc. Organisms transformed with the DNA sequences can be used
XX CC for recombinant production of the polypeptides. The new DNA sequences
XX (and their fragments) are useful as primers or probes for isolating
XX CC homologues of any of the S.aureus DNA sequences contained on the computer
XX readable medium
XX
XX Sequence 8155 BP; 3300 A; 1279 C; 1452 G; 1876 T; 0 U; 248 Other;
XX
XX Query Match 0.8%; Score 73.6; DB 2; Length 8155;
XX Best local Similarity 40.2%; Pred. No. 0.00055;
XX Matches 972; Conservative 0; Mismatches 1404; Indels 42; Gaps 5;

```

QY 3782 TTCAAGTAAACAAAGCTGTTGGTTAAATAATGATATTCGAAATGCCAATTAATTTA 3841  
Db 5651 TTAATATAAAGCACTGCAAGCTAATGAAATCAGAGCAATTTAATTAACAATTTGCAAG 5710  
QY 3842 AAGTAATGCTTTGAGTGGAGAACTCAAGTAGCAGCAGACAGAGCTTTGGAAGAGTTA 3901  
Db 5711 AGATTCCAGCTACGCCAGATGCAAGATGAAAGAAAACAGACACTGATGCTGAAGCA 5770  
QY 3902 AAGAAAGTGGAGCAAGAGAAAAGTTATCTATTGGAACTTCTGCTTCTATCAACTTAG 3961  
Db 5771 ATACTGAAATAGTAAAGCAATTCAGGCATTTTCAGCAGCACTACTAAGCAGACAGTTG 5830  
QY 3962 TGAACAATGAATTTCTGCAAAATCAGAAATTAATACGTAGCAGAGAAATCTGAAGCC 4021  
Db 5831 ATGAAGCTAAACCAATGCGAAGCAGCATTTAATGCCGTAAACACCAAAAGTTTGAAGA 5890  
QY 4022 AAAAAATGGATGTTGATGCTCACTGCTTATCAGCCGACACCCCAAGTGCAGAGAGCTTTA 4081  
Db 5891 AACAGCGCTAAAGATGAATTTGATCAATTACAAAGCAACGCAACCAATGTTATCAATA 5950  
QY 4082 ATTACAAAGCTGAAAGTCAAAATGGAACGTGAGGGGCTACTGTCAGTGGCCAATTAA 4141  
Db 5951 ATGATCAGAACGCTACACACAGAGAAAAGAAAGCAGCTATTCAACAATTAGCAACGAG 6010  
QY 4142 ACAACAAGTAATGCTTCTATTAGTGGTGGAGATATATCTAACGTTAATCAGCGGAG 4201  
Db 6011 TTACAGACGCCAAAATATAATTACAGCTGCAACTGATGATATGTTGATGATCAGGGGA 6070  
QY 4202 CAAAAGCTCTTTAGCAACCACTCAAGTACCTGTCAGTGCAGCAGCAGGAGGAGCAATTA 4261  
Db 6071 AAGAGCTGGAAAAGATTCAATTTCAAAGCACCAACAGCAACACGGTTAATCAATG 6130  
QY 4262 GTTCTGAGCGGGATTAGAAATTATCAAGGGGCTGTTCTGCTCAATAAGATTGCAATG 4321  
Db 6131 CTAAAAATGATTTGATCAAGCTGTGACAACTCAAAATCAGCAATTGTAATACAACTG 6190  
QY 4322 ACGTGAAGCTAGCGTTGATAAATCTTCATCGAAGAGCTAATGAAATCAATGTCATG 4381  
Db 6191 GTGTACAACTGAAGAGAAAATGCAAGAAAAGATTTAGTTTAAAGCTAAAGAAAAG 6250  
QY 4382 CCAAGATGTCAAGAAAGTTCTGATCTTAGCAAAAAGATATCAGGCTTTACTAAATGAA 4441  
Db 6251 CGTATCAAGATATCTTAAATGCACAACCAACTAAGATGTTACGAAAATTAAGTCAAG 6310  
QY 4442 AAGATAAAAAATATTAGAAAGATCGTGTATTATACGACTGGAAATGTTATTATAGA 4501  
Db 6311 CAGTTGCTGATATTCAGAGTATTACTGAGATACAACTTAAGATGTTG-----CGA 6364  
QY 4502 AAGAACAACTAGAAAAAGCAAGAAAAAGAAAGAGCGGTCATTTGTAATGCTGCTTAT 4561  
Db 6365 AAGATGAATTTAGCACAAAAGCAACAAAGAAAGCGTTATTGCACAAACTGCAGATG 6424  
QY 4562 CGGTTGCTGGAACGATTAATCCGCTGAGAGATAGCTATTGCACTCAATCTGTTAAA 4621  
Db 6425 CGACTACTGAGAAAAAGAAACAAGCAAAATCAACAAGTAGACGCAATTTAACACAAGGTA 6484  
QY 4622 ATTAATTTAAAGCAGAAATTGAGTGAAGAGTAAGAGACCGAGAGAGATTAATAATTCAG 4681  
Db 6485 ATCAAAATTTGAAATGCAACGTCATCGATGATGTAACCTGCAAAAGTAATTCGA 6544  
QY 4682 CGAAACATGTAATATGAGGCAAAATCATCTACTGTTGTTGTAATCGGCTTTGAGC 4741  
Db 6545 TTCAAGCAATTGACCCCAATTCAGAGATCAACAGATGTTAAAAAGAAATGCAAGAGC- 6603  
QY 4742 TTGCTATGACAGAAAGATGCTTTTTCAGAGATGGATCTTGAGCATGGAAGACTTATCA 4801  
Db 6604 TTGCTAACTGAAATGCAAAAATAAATTAACTAATAATTAATGAGACT----- 6657  
QY 4802 ATGACACGATTTGCAAGGATGATTAAGAGAAATTTCTGCTGATTCCTTAAATGGAACG 4861  
Db 6658 -----ACTAATGAAAAAAAGGTAAAGATTTAGACAGCTTGAACACACATTAAGAGAG 6712  
QY 4862 CAAATATTCATCTTGGGGGTAATGTTGCGGAACAATGCGGTTCTCTTTCTACGG 4921

Db 6713 GTTAAATNN 6772  
QY 4922 CGGTAGAGAGCTGCTTTTGCAGATACTCTTCATTAATPAAAACCTGCTTTGATTACAG 4981  
Db 6773 NNNNNNNNAAAGTTCAACACTTCTATGCAATCTCTGTTAAAGAAACAGAGTAAAAAG 6832  
QY 4982 GAACGAAGTAAATCTTTTATGAGGAAAGAAATACAAAAGTCAATGTACAACTTTGAATG 5041  
Db 6833 AATTAGATCAAGCTGACGCTATTAAGAAAAACAAATAGAAACAAACCAATGATGAC 6892  
QY 5042 ATTCTCATTTTCAAAAGCTTTCTGCTGAGGCGCTGCAGATTTAAGCAGGCTGGAATCG 5101  
Db 6893 AACCAAGAAATTAATGATGCCAAAAAGAAAGTTGATACGAATTAATCAACGAAAAAGTA 6952  
QY 5102 GAGGAATGATCTGTCAATGCTGTTCTGATGAAAGGAAAGCTTTAGTGTGTTCTG 5161  
Db 6953 ATGTGATCAATTCATCAACAAATGAATATGTTGATATGACAGTTAAAGAGAAAAAGCTA 7012  
QY 5162 AGTTGAAGAGTAAAGTTCTTCAATGTAGATGCAAAAAGATCAAAAAACAATTAATACAA 5221  
Db 7013 AAATTTAATGCAATTAAAAATTTAGTGAATACAAAAAGATGCTTTAGCTAAATTTGAAG 7072  
QY 5222 TTGCCGGAATGCAATGAGAGAAAAGCGCTGAGATTGAGCAACAGTTGCTATACAA 5281  
Db 7073 ATGCTATTAATGCTAAAGTAACGAAGCGGATTACTTAACGATCGACTTCAAGTGA 7132  
QY 5282 ATATTGGAACAATCAGTTATAGCTATTGTTAAAAACAAGTAAATTTCAACGCGGATG 5341  
Db 7133 TTGCTGAAGCGAAACAAAACCTTGCTGAATTTAAACAACTCGGATCAAAATGTTAATC 7192  
QY 5342 ATCAAGATAGAAAAAATATCATGTGACTGCAAAAAGATTACTATGACCAATACTATAG 5401  
Db 7193 AAGCTACTCTTAAAGATGACATTGAAGTTCAATTCATTAAGACTTAGATTAATTAACG 7252  
QY 5402 CAGTGGAGTTGAGAGCAAAAGAGAGCTGTGCAAGAGACTTTCGCAAGTACTACT 5461  
Db 7253 ATTAACAATTTCCAAACAGGTAAAAAAGATCAGCTAACACAGATTTATATGCTTATGAG 7312  
QY 5462 TGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAAGAGAG 5521  
Db 7313 ATCAGAAAGAAAAAATAATTTACAGCTGACACTAATGCAACAAAGATGAAAAAGCAACAG 7372  
QY 5522 AAAATTAATGAAATAAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATCAGTCAAG 5581  
Db 7373 CAATTAAGCAAGTTGACCAAAATG---TTCAAATGCTATTGAAAAGCTTAATATGTTG 7429  
QY 5582 TGTTCACAAATGCGACAGTGTCTTCCGAGCAAGTGCACAACTGCAGTAGAGCTGAG 5641  
Db 7430 TGTAAATGATGAGAGTTGATGATGATTAACACAAAGTAAAGCAGCAATGATGCTATTC 7489  
QY 5642 TAGCAGTTAATAAATTAACAAAAATCTTTCGACATATTAATAATATGACTCAAAATG 5701  
Db 7490 AAGTGAATGCTACGTTTAAACCTTAAGGAAACCAAGCTATTGAAGTTAAACAGAAAGTA 7549  
QY 5702 TAGGAATGCTTGTGTAAGA-----CAATCTCATTCATCTA 5740  
Db 7550 CGAAAGATCTATTGATCAAGAGTACCAAGTTAATCTGTAAGAAAAAATCAAGAGATTAG 7609  
QY 5741 TTAATAACATTTGGAATTTGAGAGCTGGAAGTTGAGCTGAGAGAGCTGAGAGTCAAGTTTTC 5800  
Db 7610 CAATGATTTAAACAAATTAACAGATCAAGCTAAACAAGATTAATCTGATGCAACCAACTG 7669  
QY 5801 TAGCAGTGAATAGATTGTAATAATAACGATAGCAAGATTAATCATGCAAAAATCACTG 5860  
Db 7670 CTGAAGTTGAAAAAGCAAAAGCTCAAGGACTTGAAGCACTTGAATTAACATTCAAATCGACT 7729  
QY 5861 CGAAGGAAATGTGGAGTTATTACAGAGTGTATGCGGTAAATGCTTAATTAATGCAAGAA 5920  
Db 7730 CAACAGAAAAACA AAAAGCTATCGAAGATTTAAGAACTGCACTGACCAAGATTGAAGAG 7789  
QY 5921 CAGTGTCTGAGTGGCCGCTGACAGAAATAGGAGCTCAACAGAGTGAATGAATTAACAG 5980



4439 CGAAGAACACTAGAAAAAGCAAAAAAGAGAGCGGTCACTTGTAAATGCTGCTT 4558  
 4319 CAAAAGATGAATATAGCAACAAAAGCAAAAGCAAAAGCGCTTATGTGACAAACTGCGAG 4378  
 4559 TATCGGTGCTGGAACGGAATAATCCGCTGAGAGAGTACTTTCAGTCAATACTGTGTA 4618  
 4379 ATGCACTACTAAGAAAAAGAAAGCAAGCAAAATCAACAGTAGAACACCAATTAAACAAG 4438  
 4619 AAAATTAATTTTAAAGCAGAAATTGAGTGAAGCAATTAAGAAACCGAGAGATTAATTC 4678  
 4439 GTTAATCAAAATATTTGAAAATGCAAGCTCAATGATGATTAACACTGCAAAAAGTAATG 4498  
 4679 ATGCAAAACATGTAAATGTGAGCGCAAAATCACTACTGTTGTGTGAATGCGGCTTCTG 4738  
 4499 CAATTCAAGCAATTGACCAATTTCAAGCATCAACAGATGTTAAAGCAATGCAAGAGC-A 4557  
 4739 GACTTGCTATCAGCAAAAGTGCCTTTTTCAGAAATGGATCTGAGCATGCGCAAGCTTAT 4798  
 4558 GAATTGCTAAGTGAATGCAAAATTAATTAATGTAATGTAATGTAATGTAATGTAATGTA 4614  
 4799 CAATATGACAGATTCGAAAGTGTGATTAAGAAAGAAATTTCTGCTGATTCCTTAATGTGA 4858  
 4615 -----ACTATATGAAGAAAAAGTACGATATTTGACCAAGTTAGAGCAGATATGAAAG 4666  
 4859 ACGCAATTAATTCATTTCTTGGGTGAATGTTGCGGAAACCAATGCCGCTTCTTCTTA 4918  
 4667 AAGTTTAAATTAATTAATGAGCAACTACTACAGGTATGTAATGTAATGTAATGTAATGTA 4726  
 4919 CCGCGGTAGAGCTGCTTTTGGGAATTAATCTTCAATAATTAATTAATCTGCTTGTGATTA 4978  
 4727 CAGCAGTCAAAAAGTTCAACCACTTCATGCAAACTCTGTGAAGAAACAGAGGTAAAA 4786  
 4979 CAGAAACGAAGGTAAATCTTTTAGTGAAGAAATACAAAGTCAATGTACAAGCTTGA 5038  
 4787 AAGATTTAGATCAAGCTGACGTGATTAAGAAAACAAATTAAGAACAAACCAATGAT 4846  
 5039 ATGATTTCTATTAACAAAGCTTTCTGTGAGCGCTGCAAGTAATTAAGAGCTGGA 5098  
 4847 CACAACAAAGAAATTAATGATGCAAAAACAGAAAGTTGATCTGAATTAATTAATCAAGGAAA 4906  
 5099 TCGAGAGAAATGATCTGTCATTCGCTGTTCTGATGAAGCAAGCTTATGATGAT 5158  
 4907 CAATGTGATCAATCATCAACAAAGTAATGTTGATATGATCAATTAAGAAAGAAAG 4966  
 5159 CTGAGTTTGAAGAGTAAGTTCTTTCAATGTAGATGCAAAAGATCAAAATCAATAATA 5218  
 4967 CTAAATTAATGAGTAAACATTTAGTGAATCAAAAAGATGCTTAAAGTAAATG 5026  
 5219 CAATTGCCGAAATGCAAAATGAGAAAGCGGCTGAGTGAAGCAACAGTTGCTCAT 5278  
 5027 AAGATGCAATATATGCTAAAGTAAAGCAAGCGGATTAATCTTAACGATCGACTCAAGTG 5086  
 5279 CAATATTTGGAACCAATCAATGTTATGCTATTTGTAATAAAACGTAATAATTAACAGCGGA 5338  
 5087 AATTTGCGAAGCAAAACAAACCTTGCAATTAACAAACCTGCGATCAAAATGTTA 5146  
 5339 ATGATCAAGATAGAAAAATATCAATGTGACTGCAAAAGATTAATGATGACCAATACTA 5398  
 5147 ATCAAGCTACTCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5206  
 5399 TAGAGTGGAGTGGAGAGCAAAAGAGCCTGTGTGCAAGAGCTTTCGCAAGTACTA 5458  
 5207 ACGATTACACCAATTCACAGAGTAAAGAAATCAAGCTCAACAGATTAATGCTTAATG 5266  
 5459 CTTGAATTAAGACAGTTCTTCTCATGTTGATCAACCTGATTTGCAAAAGTTTAGAG 5518  
 5267 CAGATCGAAGAAAAATTAATTTTCACTGACACTTAATCAACACAGTGAAGAAAGCAAC 5326  
 5519 AAGAAATTAATGGAATTAAGAAAGCAAAATGTTAATGTTAGTGAAGAAATTAAGAGT 5578  
 5327 AAGCAATTAAGCAAGTTGACCAAAATGTTCA---AATGCTATTTGAAAAATTAATGAT 5383  
 5579 AAGTGTCACAATGCGACAGTCTTTCCGAGCAAGTGAAGCAAGCTGCAAGTGAAGAGCTG 5638

Db 5384 GTGTGATTAATGAGTACGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5443  
 QY 5639 GAGTAGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5698  
 Db 5444 TTCAAGTAGATGCTACTGTTTAACTTAAGCAACCAAGCTTATGAAAGTAAAGCAGAG 5503  
 QY 5699 AT 5700  
 Db 5504 AT 5505

RESULT 28  
 ACF05848  
 ID ACF05848 standard; DNA; 7446 BP.

ACF05848;  
 04-DEC-2003 (first entry)  
 Methicillin resistant Staphylococcus aureus ORF SA1964-fmbB (mrp).

Vancomycin; teicoplanin; MRSA; diagnosis; SA1964; mrps; fmbB;  
 antibiotic resistance; gene; ds.

Staphylococcus aureus.

Key location/Qualifiers  
 CDS 1..7446  
 /\*tag= a

W02003062466-A2.

31-JUL-2003.

16-JAN-2003; 2003WO-GB000129.

16-JAN-2002; 2002GB-00000865.

(UNBR-) UNIV BRISTOL.

Walsh TR, Avison MB, Howe RA;

WPI; 2003-636743/60.

P-PSDB; ABR62804.

Detecting a strain of Staphylococcus aureus having reduced susceptibility  
 to vancomycin and/or teicoplanin, comprises detecting the presence of  
 mutation(s) leading to loss of function of at least one gene.

Claim 7; Fig 3; 78pp; English.

The present sequence is that of open reading frame ORF SA1964-fmbB (mrp)  
 CC 2218122-2225567 of methicillin resistant Staphylococcus aureus (MRSA)  
 CC strain N315. The ORF was identified as an mrp/fmbB gene. The invention  
 CC provides a method of detecting a strain of Staphylococcus aureus which  
 CC has reduced susceptibility to vancomycin and/or teicoplanin by detecting  
 CC the presence of one or more mutations in at least 2 open reading frames,  
 CC including the present sequence. Polypeptides, polynucleotides,  
 CC antibodies, probes, primers, kits and methods are provided for this  
 CC detection. The methods are rapid, reproducible and sensitive, and allow  
 CC for detection of strains that are fully or heterogeneously resistant

Sequence 7446 BP; 3198 A; 1312 C; 1311 G; 1625 T; 0 U; 0 Other;

Query Match 0.7%; Score 72.4; DB 10; Length 7446;  
 Best Local Similarity 41.8%; Pred. No. 0.00092;  
 Matches 804; Conservative 0; Mismatches 1091; Indels 27; Gaps 5;

QY 3782 TTCAAGTAAAGCAAGCTTGTGTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3841  
 Db 3608 TTAATTAAGCAACAGCTGACGTAAATGAAATCAAGCAATTTTAATTAACAAATTTGCAAG 3667

QY	3842	AAGTAAATGCTTTGATGAGGAACTCAAGTACACAGAGAGGTTTGAAGCGTTA	3901
Db	3668	AGATTCAAGCTACACCATGATGCAACGATATGAAAAAACAAGAGCTGATGCAAGCA	3727
QY	3902	AAGAAATGAGAGCAAGAAAAAGTATATTTGGAACTTCTGCTTATCAACTTAG	3961
Db	3728	ATACTGAAATGTTAAAGCAATTCAGCCATTTTCAGAGCAATCTTAACGCAAGTTG	3787
QY	3962	TGAACAATGAAGTTCTGCAAAATCAGAAAAATATACAGTAGAGAGAACTTGAAGCC	4021
Db	3788	ATGAAAGCAAGCTAAATGCAAGACAGCAATTAATGCGGTAAACACAAAAGTTTAA	3847
QY	4022	AAAAAATGATGTTGATGATCACTGCTTATCAAGCGGACACCCAAAGTAGACAGAGCTTAA	4081
Db	3848	AACAAGCGGCTAAAGATGAGATTGATCAATTAACAGACAAAGCAAAATGTTATCAAT	3907
QY	4082	ATTTACAGAGCTGAGAAAGTCAAAATGAACTGAGGGGCTACTGACCTGTGCCAAATTA	4141
Db	3908	ATGATCGAAGCGTACAAATGAGAAAAAAGACAGCTATTCAACAATTAGCAACGAG	3967
QY	4142	ACAAACAAGTAATATGCTTCTATTAGTGTGGAGATATTAACGTTAATCGACGAGC	4201
Db	3968	TTACAGACGGCAAAAATATATTTACAGCTGCAACTGATGATTAATGTGTGATACAGCA	4027
QY	4202	CAAAAGCTCTTTTGACCAACCTCAAGTACTGTCGACGTGACGCGGAGGAGCAATTA	4261
Db	4028	AAGACGCTGAGAAAGATTCATTTCAATTCMAAGCACACAACAGACAAACAGCAG-----TTA	4078
QY	4262	GTTCTGAGCGGGATTTGGAAATTAATACAGGGGCGTCTTCTGTCAATPAG--ATTGACA	4318
Db	4079	AATGAAATGCGAAAAATGAAATTGATTCAGCTGTGACAACTCAAAATCAGCAATTGATA	4138
QY	4319	ATGACGTGGAAGCTAGCGTTGATTAATCTTCCATCGAAGAGCTAATGAATCAATGTCA	4378
Db	4139	ATACAACTGGGGCTACAACTGAAAGAAAAATGACGAAAAAGATTAGTTTAAAAAGCTA	4198
QY	4379	TTGCCAAAGATGTCAAGAAAGTTCTGATCTAGCAAAAGAAATATCAGAGCTTACTAATG	4438
Db	4199	AAGAAAAAGCGTATCMAATATCTTAAATGCAAAACAACTAAAGATGTTACGAAATTA	4258
QY	4439	GAAAAAGATAAATAATTTAGAAGTCGTGTTAATTAATACGACTGAGAAATGTTAATATA	4498
Db	4259	AAGATCAAGCAGTTGCTATATTTCAAGTATTAATCTGAGATACAACAATTTAAGATGTTG	4318
QY	4499	CGAAGGAACAATAAGAAAAACAAGAAAAAGAAAGAGCGGTCAATTGTAATGCTGCTT	4558
Db	4319	CGAAAGATGAATTAGCACAAAAGCAAAAGCAAAAGCGCTTAATGCAAACTGCGAG	4378
QY	4559	TATCGGTTCTGGAACGGAATAATCCGTGAGAGATGACTAATTTGCACTCAATCTGTTA	4618
Db	4379	ATGCGACTACTGAGAAAAAGAAACAAGCAAAATCAACAAGTGAACGCAATTTACACAG	4438
QY	4619	AAAAATTAATTTAAAGCAGAAATTTAGTGAACCAATTAAGSAGCCGAGAGGATTAATTC	4678
Db	4439	GTAATCAAAAATTTAAAAATGCAAGTCAATGATGATGTAAACCTGCAAAAGATATATG	4498
QY	4679	ATGCAAAACATGTAATATGAGGCAAAATCATCTAATCTGTTGTTGTGATGCGGCTTCTG	4738
Db	4499	CAATTCAGCAATTTGACCAATTTCAAGCATCAACGATGTTAAAAAGAAATGCAAGAC-A	4557
QY	4739	GACTTGCTATCGCAAGAATGCTTTTTCAGAAATGGGATCTGAGACATGCAAGACTTAT	4798
Db	4558	GAAATGCTAACTGAATATGCAAAATTAATACTGAATAATCTTAATTAATGAAAGCT---	4614
QY	4799	CAAAATGACAGATTGCAAAAGTGATTAAGAAAGAAATTTCTGCGATTCCTTAATGGA	4858
Db	4615	-----ACTAATGAGAAAAAGGTAAAGATATTGGACCAAGTTAGACAGATATGAG	4666
QY	4859	ACGCAAAATATTCATTTCTTGGGGTGAATGTTGCGGGAACAATTCGCGGTTCTCTTTCTA	4918
Db	4667	AAGTTTAATTAATTAATTAATGACGAACTACAGAGATGTAATCTACTCTAAAGATA	4726
QY	4919	CGGCGGTAGAGCTGCTTTTGGCAATATATCTTTCAATATTAACCTCTGCTTGAATTA	4978

Db		4727	CAGCAGTACAAAAGTTCAACAATCTTCATGCAAAATCTGTTAAGAACCAGCAGTAAAA	4786
Oy		4979	CAGGACGAAGGGTMAATCCTTTTAGTGGAAGAATACAAAGTCAATTGACAAGCTTTGA	5038
D4b6		4787	AAGAAATTAGTCAAGCTGCAGCTGATAGAAAACAAATATGAAACAACCAATGCGAT	4846
Oy		5039	ATGATTCTCATATTACAAACGTTTCTGCTGAGGCCCTGCAGATATTAGCAGCGTGGAA	5098
Db		4847	CACAACAAGAAATTAATGATGCAAAACMAAAGTTGTATCTGAATTAAATCAAGCCGAAA	4906
Oy		5089	TCCGAGGAATGATATCTGTCAATGCTGGTTCTGATGTAACGGAAGCTTTGTTAGTATT	5158
Db		4907	CAAAATGTCGATCAATCATCAACAATTAATATGTTATATGCAATTAAAGAGAAAAG	4966
Oy		5159	CTGAGTTTGAAGAGTAAAGTTCTTTCATGTAGATGCAAAAGATCAAAAAACATAAATA	5218
Db		4967	CTAAAAATTAATGCGATTAAAAACTTTTAGTAGTACABAAAAGATGCTTTAGCTAAATTTG	5026
Oy		5219	CAATTGCGGAATGCAAAATGAGAGAAAAGCGCTGAGTTTGAGCAACAGTTCCTATA	5278
Db		5027	AAGATGCAATATPATATGCTAAAGTAAACGAAGCGGATTACTGTAAGCATGCACTTCAAGTG	5086
Oy		5279	CAAAATTTGGAAAAACATCACTTATACCTATTTGTAATAAACAGTAAATTAACAACGGCGA	5338
Db		5087	AAATTTGCTGAAGCGAAACMAAACTTCTGTAATTAAAACAAACCTGCGGATCAAAATGTTA	5146
Oy		5339	ATGATTCAGATAGAAAAAAATATCAATGTGACTGCAAAAGATTATATATGACCAATACTA	5398
Db		5147	ATCAAGCTACTCTTCAAGATGACATTTGAAGTTCAAAATTCATATAGACTTAGATTAATTA	5206
Oy		5399	TAGCAGTCGAGTTGAGAGACAAAGAGCGCTGTGCAAGAGCTTTGCAAGTACTA	5458
Db		5207	ACGATTTACAAATTCCAACAGGTAAAAAAGATCAGTACCAACAGATTTATATGCTTATG	5266
Oy		5459	CCTTGATTAAGACAGTTCTTCTCATGTGTGTCAAACTGATATATGACAAAGATTAGAGG	5518
Db		5267	CAGATCACAAAAAAATATATATTTTCAGGTGCACTAATGCAACCAAGATBAAAAGCAAC	5326
Oy		5519	AAGAAAAATATGAGAAATAGAGAAAAAGCAAAATGTTAATGTTCTGAGCAAAATACGATC	5578
Db		5327	AAGCAATTAAGCAAGTTGACCAAAATGTTCA---AATGCAATTGAAAAACATTATATATG	5383
Oy		5579	AAGTGTGCACAAATGCGACAGTGTCTTTCGAGAGCAAGTGCACAAGCTGCAATGAGAGCTG	5638
Db		5384	GTTGTGATATATGTTGATGCGTTGATGATTAACACAAGGATPAAAGCAGCAATTTATACTA	5443
Oy		5639	GAGTAGCAGTTATATAAATTACACAAAATACCTTCTGCACTATATAAAAAATGTACTCAA	5698
Db		5444	TTCAAGTAGATGCTACTGTTAAACCTTAAGGAAACCAAGCTATTTGAACTTAAGCAGAG	5503
Oy		5699	AT 5700	
Db		5504	AT 5505	
 RESULT 29 ADU74046 standard; DNA; 7446 BP.				
ID	ADU74046			
XX	ADU74046;			
XX	10-FEB-2005	(first entry)		
DE	Staphylococcus aureus Mtp homologue coding sequence.			
XX	ANTIBIOTIC-RESISTANCE; VRSA infection; Mtp homologue; gene; ds.			
OS	Staphylococcus aureus.			
FH	Key	Location/Qualifiers		
FT	CDS	1..7446	/tag= a	

/product= "Mrp homologue"

MO200409444-A2.  
18-NOV-2004.  
07-MAY-2004; 2004WO-GB001985.  
09-MAY-2003; 2003GB-00010794.  
(UVR-) UNIV BRISTOL.  
Walsh TR, Avison MB, Howe RA, Wootton M;  
WPI; 2004-821685/81.  
Detecting a strain of *Staphylococcus aureus* with reduced susceptibility to vancomycin or telicoplanin by relating the detected expression level of *fmfB*/*mrp* or homolog within the strain to susceptibility to vancomycin or telicoplanin.  
Claim 2; SEQ ID NO 17; 102bp; English.  
The present sequence is the coding sequence for an Mrp-like protein of a vancomycin-sensitive *Staphylococcus aureus* strain, denoted SA1964-fmfB (mrp) 2218122-2225567. The encoded Mrp homologue protein may have a similar function to *S. aureus* Mrp/*fmfB* proteins. Expression of the *mrp* gene is down-regulated in *S. aureus* strains with a decreased susceptibility to vancomycin. The invention provides a method for detecting a strain of *S. aureus* which has reduced susceptibility to vancomycin and/or telicoplanin. The method involves detecting the level of expression of an *fmfB*/*mrp* gene or an *mrp*-homologue gene (including the present sequence) or a variant within the strain and relating that to the susceptibility or resistance of the strain to vancomycin and/or telicoplanin. Detection of expression is performed using an immunological method to detect the gene product or a molecular method such as RT-PCR to detect RNA constructs. The method is rapid, reproducible and sensitive and allows, for example, discrimination of vancomycin intermediate *S. aureus* (VISA) and heterogeneous vancomycin intermediate *S. aureus* (hVISA) from vancomycin/telicoplanin sensitive *S. aureus* (VSSA).  
Sequence 7446 BP; 3198 A; 1312 C; 1311 G; 1625 T; 0 U; 0 Other;  
Query Match 0.7%; Score 72.4; DB 13; Length 7446;  
Best Local Similarity 41.8%; Pred. No. 0.0092;  
Matches 804; Conservative 0; Mismatches 1091; Indels 27; Gaps 5;

3968 TTACGACGCGCAAAATAATATTAACAGCTGCACCTGATGATATGTGTATGATACAGCGA 4027  
4202 CAAAAGCTCTTTAGCAACCACTCAAGTACTGCTGCAGTACGACGCGAGGACAAATTA 4261  
4028 AAGACGCTGGAAAATAATTCATTCAGACACACACGACACGACG-----TTA 4078  
4262 GTTCTGAGCGCGATTAGAAATTATCAAGGGGCTGTTTCTGTCAATAG--ATTGACA 4318  
4079 AATCGAATGCGAAAAATGAAAGTTGATCAAGCTGTGACAACTCAAAATCAGCAATGATTA 4138  
4319 ATGACGTGGAAGCTTAGCTGATTAATCTTCATCGAAGAGCTTAATGAAATCAATGCA 4378  
4139 ATCAACTGGCGCTACCACTGAGAGAAAAATGACGCAAAAAGATTATTTAAAGCTA 4198  
4379 TTGCAAAAGATGTCAAAAGAAAGTTCTGATCTTAGCAAAAAGAAATACAGCTTTACTAAATG 4438  
4199 AAGAAAAGCGTATCAAGATATCTTAATGACAAACAACTAATGATGTTACGCAATTA 4258  
4439 GAAAAATAAAAAATTTAGAAAGATCGGTATTAATACACTGAAATGTTATTA 4438  
4259 AAGATCAAGCAGTTGCTGATATTCAGGTATTTACAGATACAACTTAAGATGTTG 4318  
4439 GAAAGAACAACTAGAAAAAGCAAAAGAAAAGAGCGGCTATTGTAAATGCTGTT 4558  
4319 CGAAAGATGAATTAGCAACAAAGCAAAAGCAAAAGCGCTTATTTGCAAACTGCG 4378  
4559 TATCGGTGCTGGAACGGATTAATCCGCTGAGAGGTAGCTAATGCACTGTAATCTGTTA 4618  
4379 ATGCACTACTGAAAGAAAAGAACAGCAATCAACAGTACAGTACACATTAACACAG 4438  
4619 AAAATAAATTTAAAGCAAAATTAAGTGAAGCAATTAAGAAACCGAGAGATTAATTC 4678  
4439 GTAATCAAAATATTGAAATAATGACAGTCAATCGATGATGTAACATGCAAAAGTAAATG 4498  
4679 ATGCAAAACATGTAATGTGAGCGCAAAATATCTACTGTTGTGTAATCGGCTTCTG 4738  
4439 CAATTCAAGCAATGACCAATTCAGAGATTAACAGATGTTAAACGAATGCAAGAG-A 4557  
4739 GACTTGCTATCAGCAAAAGATCTTTTACAGAAATGGGATCTGAGCATGCGAAGCTTAT 4798  
4558 GAATTCCTACTGAAATGCAAAATTAATTAATCTTAATTAATGAACT--- 4614  
4799 CAAATGACACGATTCGAAAGGTGATTAAGGAAGAAATTTCTGCTGATTCCTTAATGTA 4858  
4615 -----ACTAATGAAGAAAAGGTAAAGTAAAGTATGACAGTAAAGCAGACATATGAAG 4666  
4859 ACGCAATTAATTCATCTTGGGTGAATGTTGGCGGAACCATTTGCCGTTCTCTTTCTA 4918  
4667 AAGGTTTAATTAATTAATTAATGACCAACTACTACAGGTGATGTAATCTGCTAAAGATA 4726  
4919 CGGCGTAGAGCTGCTTTTGGCAATTAATCTTCAATTAATAAAGCTGTGCTTTGATTA 4978  
4727 CAGCAGTACAAAAGTTCAACAATCTTCATGCAATCTGTTTAAAGAACAGAGTTAAA 4786  
4979 CAGGAACGAAGTAAATCTTTTAGTGAAGAAATACAAAAGTCAATGTACAACTTTGA 5038  
4787 AAGAAATGATCAAGCTGACAGTGAAGAAAACAAATGAAACAAACCAATATGAT 4846  
5039 ATGATTCATATTAACAACGTTTCTGCTGAGCGCTGCAAGTATTAAGCAGCGTGA 5098  
4847 CACAACAAGAAATTAAGTATCAAAAACAAAGAAAGTTGATACGAATTAATCAACGAAA 4906  
5099 TCGAGGAATGTTATCTGTCAATGCTGTTCTGATGAAAGGAAAGCTTTAGTTAGTAT 5158  
4907 CAAATGTCATCATCATCAAAATGAATATGTTGATTAATGCAATTAAGAGAAAG 4966  
5159 CTGAGTTTGAAGGATAGTTCTTTCAATGTAGATGCAAAAAGATCAAAAACAAATTAATA 5218  
4967 CTAATTAATTAATGCAAGTTAAACATTTAGTGAATCAAAAAGATGCTTTACTAAATATG 5026  
5219 CAATTTGCCGGAATGCAAAATGAGAAAAGCGCTGAGATTGGAACAACTGCTCATTA 5278  
5027 AAGATGCAATTAATGCTTAAATGAACGAAGCGATTAATCTTAAGCATGCACTTCAAGTG 5086

Oy	5279	CAAAATATTTGGAAAAACAATCAGTTATAGCTATTGTAAAAAACAAGTAAAAATTACAAAGCGCA	5338
Db	5087	AAATTTGCTGAAGCGAAACAAAACTTGCTGATTTAAAAACAACCTGCGATCCAATATGTTA	5146
Oy	5339	ATGATCAAGATAGAAAAAATATCAATGTCAGTCCGAAAGAATTACTATATGACCAATCTACTA	5398
Db	5147	ATCAAGCTACTCTTTAAAGATGACATTGAAGTTCAAAATTCAATATGACTTGAATATATTTA	5206
Oy	5399	TAGCAGTCGAGTTGGAGAGCAAAAGAGCCTCTGTGCAGAGAGCTTTCTGCAAGTACTA	5458
Db	5207	ACGATTCACACATTTCCAACAGATGAAAAAGATCAGTACCAACGATTTATATATGCTTATG	5266
Oy	5459	CCTTGAATTAAGACAGTTTCTTCTCATGTTGATCAACATGATATTTGACAAAGATTATAGG	5518
Db	5287	CAGATCAGAGAAAAATATATTTTTCAGCTGCACTAATGCAACACAGATGAAAAAGCAAC	5326
Oy	5519	AAGAAAAATATATGAAATTAAGAAAAAGCCAAATGTTAATGTTCTAGCTGAAAAATACGAGTC	5578
Db	5327	AAGCATTTAACAAGTTGATGCCAAATATGTTCA--AATGCAATTAGAAAAACATTAATATATG	5383
Oy	5579	AAGTGTGTACAAATATGCGCAGTGCCTTTCCGAGCAAGTGCACAGCTGCAGTATAGACTG	5638
Db	5384	GTTGTGATTAATATGTTGACGTTATGATGATTTAAACAAAGTAAAGCAGCAATTTATATCTA	5443
Oy	5639	GAGTAGCAGTTAATTAATTAATTTACACAAATACTTGTGCATATATAAAAATAGTACTCAAA	5698
Db	5444	TTCAAGTAGATGCTACTGTTTAAACCTTAAGGAAACCAAGCTATATTGAAGCTTAAGCAGAG	5503
Oy	5699	AT 5700	
Db	5504	AT 5505	

CC	RESULT 30
XX	AAA70212
ID	AAA70212 standard; DNA; 1998 BP.
XX	
AC	AAA70212;
XX	
DT	07-NOV-2000 (first entry)
XX	
DE	Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:345.
XX	
KM	Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine; antimalarial; malaria; protozoacide; infection; insecticide; ds.
XX	
OS	Plasmodium falciparum.
XX	
PN	WO200025728-A2.
XX	
PD	11-MAY-2000.
XX	
PF	05-NOV-1999; 99WO-US026796.
XX	
PR	05-NOV-1998; 98US-0107131P.
XX	
PA	(HOFF/) HOFFMAN S.
PA	(CARU/) CARUCCI D.
PA	(GARD/) GARDNER M.
PA	(VENT/) VENTER J C.
XX	-
PI	Hoffman S, Carucci D, Gardner M, Venter JC;
XX	
DR	WPI; 2000-365347/31.
XX	
PT	Proteins encoded by chromosome 2 of the human malarial parasite, Plasmodium falciparum, useful as antimalarial vaccines and in the diagnosis of P.falciparum infection.
XX	
PS	Disclosure; Page 537; 577pp; English.
XX	
CC	The present invention describes proteins and their fragments (I) encoded

by chromosome 2 of the human malarial parasite, *Plasmodium falciparum*. Also described are: (I) nucleotide sequences (II) encoding (I); and (2) vaccines against *P. falciparum* infection comprising (I) or (II), (I) and (II) are useful for the development of vaccines against *P. falciparum* infection. (I) and polyclonal antisera or a monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with *P. falciparum*. Furthermore, (I) (especially when they are rifins or secreted or membrane proteins) can aid the identification of drugs to treat or prevent *P. falciparum* infection, or they can be used to identify drug resistance in *P. falciparum*. Sequencing of the *Plasmodium* chromosome 2 and the subsequent identification of parasite proteins encoded by it will help to expand our understanding of parasite biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AA70078 to AA70287 and AB8144 to AB81352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification

Sequence 1998 BP; 1127 A; 141 C; 338 G; 392 T; 0 U; 0 Other;

Query Match 0.7%; Score 68.8; DB 3; Length 1998;

Best Local Similarity 42.5%; Pred. No. 0.0031;

Matches 434; Conservative 0; Mismatches 577; Indels 9; Gaps 1.

QY 5130 TGATGAACCGAAGCTTCTAGTGAATCTGAGTTTGAAGAGTAAAGTCTTTCATAT 5188

Db 816 TGATGAAGATGTAAGTGTATGATGATGTTGTAAGAAGATAAATAATGAATGATGATTTGA 875

QY 5190 AGATGCAAAAGATCAAAAAACAAATAAATACAAATTCGCCGAAATGCAAAATGAGAGAAAAAGC 52443

Db 876 TGATGATATAAAGCAACAGATTAACAACATTTGGAAGAGAGGAAATGAATTAATTGA 935

5250 GGCTGGAGTTGG-----AGCAACAGTTGCTCATACAATATTTGGAACAACATCAGT 5300

Db 936 AAAGCAATTAGTGATAGAAAAGATGCTAAAATAAGACTACTAAGAGGAAAAAG 995

5301 TATAGCTATTGTAAAAACAGTAAATTACACGGCAATGATCAAGATGAAAAAATAT 5360

996 TAAAGATTTCTGAGAGGAAAAAGTTAAAGATTTAGAGAGCAAAAAAGTTAAAGATTAAGA 1055

5361 CAATGTGACTGCAAAAGATTATACATATGACCAATACTATATGACAGTCCGAGTTTGAGAGCC 5422

Ph 1056 GAGCGAATAAAGTAAAGAGAGAAAGCTAAAGATTAAGGAGAGGAAAAAG 1111

[illegible][illegible][illegible][illegible]

**Dd** 1176 TACTCGGAGGAAAAAGAAAAGATATAAGACAAAGAAAAAAGCAGAACAATAAGCA 123

5541 AAGGCAATGTTAATGTTCTAGCTGAAAVACGAGTCAAGTGGTCACTAATGCGACTG 5800

Db 1236 AAATTAAAGCATAAACCAATGATGAAAAAAGAAAGACGCTAACGAAAAAGAAAAAA 129

5601 GCTTCCGAGCAAGTGCACAGCTGCAGTACGAGTATATAAAATTAC 566

Db 1296 TGATAACAGATATACATGATGATGATGAAATGATATGGAAGAATAGAGA 135

Qy 5661 ACAAATACCTTCTGCACATATAAAAAATAGTACTCAAAATGTACGAAATGCTTGGTAA 5722

Db 1356 AAATGATGACGAAGATGAGATGAAGACATGGAAAAACAAAAAAAAAAAAAGG 1411

5721 AAGCAATCTCATTCATATTAACAATTGGAATTGAGCTGGAGCTGGAGG 578

Db  
1416 AAAAATGGAATGAAATCGAAATGAAAAGTGAATAAGAAAATGGAATGAAAAATGGAATGAAA 1477

5781 AGCTGAGTGCAGTTCTGTAGCAGTATAGATTGTAATAATATACGATAGCAGATT 584





CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX

Sequence 4997 BP; 2206 A; 790 C; 905 G; 1096 T; 0 U; 0 Other;

Query Match 0.7%; Score 68.2; DB 8; Length 4997;

Beet Local Similarity 41.0%; Pred. No. 0.0054;

Matches 933; Conservative 0; Mismatches 1318; Indels 24; Gaps 5;

QY	5215	AATACAAATGCGGAAATGCAATATGAGAGAAAAGCGCTGAGTTGAGCAACAGTTGCT	5274
DB	1969	AAAGCAGCAGCAAGAAAAGCGATTATATGATGACGAACTGCTCAAAAAAGAACGTAAT	2028
QY	5275	CATACAAATATTGAAAAACATCACTTATAGCTATTGTAAAAACAGTAAATTTCAACG	5334
DB	2029	AATAGCTCAGATGCAACACAAAGAAAAGATGACGAATATGCAAAAGTTGATCAGCT	2088
QY	5335	GCGATGATCAAGATGAAAAAATATCAATGATGCAAAAAGATTATATATGCAAT	5394
DB	2089	GTAACAGCTGCAAAACAAAGCAATTTACACAGCAACAAACAAATGATAT- -GTAGCCAA	2146
QY	5395	ACTATGACAGTCGAGTTGAGAGCAAAAAGAGCCTCTGTCAGAGAGCTTTCGCAAGT	5454
DB	2147	AACAAATAGCGGTACTTCTGACTATTTACTGTTATCAACACAG- -AAGTTATCTAAAAACA	2205
QY	5455	ACTACCTTGAATTAAGACAGTTCTTCTCATGTTGATCAATGATATTTGACAAAGTTTA	5514
DB	2206	GCAGCAAGAAAAAGCAATTTGATGATGAAGTCGTTGAAGAAAGCAATTTGATCTTT	2265
QY	5515	GAGGAGAAATTAATGAAATTAAGAAAAAGCAATTTTATGTTCTTCTGCTGAAATTCG	5574
DB	2266	GCCGATGCTACAGATGAAGAAAAACAGCAGCTTAAAGATTAAGTTGATGCTGAACTTCC	2325
QY	5575	AGTCAAGTGTGCACAAATGCGCAGTGTCTTCGAGAGCAAGTGAACAAGCTGCAATAGA	5634
DB	2326	AAAGCAAAAGCGCAATTTGATCACTACTACAAACAAAGATGTAGCAAGCAAAATCA	2385
QY	5635	GCTGAGTAGCAGTTTAATTAATTAACAAATATCTTGCACATATTAATAATAGTACT	5694
DB	2386	TCTGTGTAAACAATATTGAAGAAATACAACTGATGATTAAGAAATTCGACAGAAAA	2445
QY	5695	CAAAATGTACGAATGCTTTGTTAAAGCAAAATTCATTCATCTTATTTAAACAAATTGGA	5754
DB	2446	CAAGCAATATGATGATGCAAAATGCTTAAAAAGACGAATTTGCAATTAATAGTATGCA	2505

QY	5755	ATTGAGCTGAGATTGAGAGCTGAGAGCTGAGTGAACAGTTCTGTAGCAGTAATAG	5814
DB	2506	ACTCAAGAAAGAAAGACGTAGCAAAAGCAAAAGTATGTCGAAGCTGCTTAAAGCTTAA	2565
QY	5815	ATTGTAATATATAGATATGCAATTAATATCATGCAAAAAATCATCTGCAAGGAAATGTC	5874
DB	2566	ACACGCAATTGATTAAGCAACAAAGATATGATGTCACAAAGAGACTTATGGCAAC	2625
QY	5875	GGAATTTTACAGAGCTGATGCGGTATTTGCTAATTAATGAGAAACAGTGTCTGAGTGG	5934
DB	2626	AATTCATTTTCAATTAATTCAGCTGATATTTGTTAAAAAGACGACGAAAGAAAGCAATT	2685
QY	5935	GCCCGTACAGCAATAGAGCCTCAACAGTGTGAATGAATTAAGATTTCAAAAG- -	5992
DB	2686	GATATGACAGCAACGCTTAAGAAACAAAGAAATGATCAACCCCAACAGTCAACAAGAA	2745
QY	5993	----CATATGTAAAAAGTTCTTACAGTATGCTTAAAGAAACAGATATTTATTTACT	6048
DB	2746	GAAGAAAGCGCAGCAAAAAGCTTAAGTATGCTGAAGTACAAAAGCGAAAGATGCAATT	2805
QY	6049	ACTCAAGGCAAGTATGAATTAAGTGTATTAAGTATTTCAAAAATCTTAAATTAACGAA	6108
DB	2806	ACTCAAGCAATTTCAAAATTAATGATGTATGATCAAAATTTCTGAAACTGCTTCAATA	2865
QY	6109	GACTTATCACAAAAAGAAAAAATTAAGTAAATTAAGAAAGATTTTACCAATATGTTCA	6168
DB	2866	GCAGATTTTCAAGTATGATGACAGTTTAAAAAGCAAAATGCTAAACAGCAATGATGATGA	2925
QY	6169	ACTCATACTTTAAATCTTTATTTGGCAATGCGCTGCTGTTGAGACAAAGCGAGTGGCA	6228
DB	2926	GCAGCTAAGAAAAATGAAATTTGACCAACATCCAAACACTACACAAAGAAAGATGCA	2985
QY	6229	GGAATGTTAAT- -CAACAGTTTATGAGAAACAGAAAGCTTTGAGAAATTTCT	6285
DB	2986	GCAAAAGCAAAAGTATGAAAGAGTACTTAAACCAAAACGAGCAATTTGATTCATCGACT	3045
QY	6286	ATATTAATGCAAAACATTTATTTCTGTAAATCAGAGATTTACAGAAATTCATCGAGTA	6345
DB	3046	ACGAATGACCGGTATGATTCAGAGCAAAAGACTCTGTGTATTAATTAATTAATTTCAA	3105
QY	6346	GTAGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	6405
DB	3106	CCAAATCAATTAAGAAAAATTTAAAGCAAGCAACGATTTGATGATCTGTTTACAGCCAA	3165
QY	6406	ATTAAGAAATATACCAAGCAAGAGTTGAAAAATCTAAATGCTGATGAAGTTTGGGA	6465
DB	3166	AAGAACGCTATGATCAAGATGTAATCTTACAAAGAAAGAAAAAGATGCTTAAAGCA	3225
QY	6466	GAAAGCTGAAATTTACAGCAGATTTCTTAAGCAAGAAATTTCTTTTGGAGTCGAGATC	6525
DB	3226	AAAGTATGAAAGAAAGTAACTTAAAGCGAAACAACTTATATGATCAAGCTATGATATGAT	3285
QY	6526	GCAGCAGCGGGGTAGAGCCGAGTGGCAGAAACGTTTCCGTAATCAATTTGACAGA	6585
DB	3286	GCTGTACAGCAAACTCGAGATTAATGAATTAATCTGTTAAACAAATTTCAACTGAAAGCA	3345
QY	6586	AAGACGAAATGATGTGGAAGAAAGCAAAAGTTTGTGTAAAAAAGCTGAGATTTACGCA	6645
DB	3346	ATCAAAAAAGCTGATGAAACAAAGCAATCGATATGCTGTACAGCTTAAAAAAGCAAGCA	3405
QY	6646	AAAGCTTATAGTTCTGTTGCAATTTGAAATGCGCAGTGGAGTGGCTGCAAAAGAGACT	6705
DB	3406	ATATATCAAGTTTCTGATGCTTACCAAGAAAGAAATGACGAAAGCTTAAAGTTGAC	3465
QY	6706	GGAATGAGCAGCAGTGGCAGTTACCAAGATGAATCAACACAGAGCAAGAGTGA	6765
DB	3466	GAAAGATTAACAAAGCTAAGTCAAGCATGCAATTAAGCAACAAATTAATGATGTTGAC	3525
QY	6766	AATTTTAAATTTATGACTCGAAACAAAGTATGATGATTAATGAGCAAAATGAGATTAATCA	6825
DB	3526	CAAGCAAAAGCAAGTGAATCTACATTAATTTCTTCAATTAAGAACGCA- - - - -A	3573
QY	6826	GGTACTGGAATCGGTTTACCGGAGCTGGAATTTCTTGACCGGAGATATCTGAGTGGTT	6885



XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;  
 KM tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;  
 KM cytosine methylation; de.  
 OS Homo sapiens.  
 XX MO200168912-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 15-MAR-2001; 2001WO-EP002955.  
 XX  
 PR 15-MAR-2000; 2000DE-01013847.  
 PR 06-APR-2000; 2000DE-01019058.  
 PR 07-APR-2000; 2000DE-01019173.  
 PR 30-JUN-2000; 2000DE-01032529.  
 PR 01-SEP-2000; 2000DE-01043826.  
 XX  
 PA (EP1G-) EP1GENOMICS AG.  
 XX  
 PI Olek A, Piepenbrock C, Berlin K;  
 DR WPI; 2001-602752/68.  
 XX  
 XX Fragments of chemically modified genes associated with tumor suppressor  
 PT genes and oncogenes, useful in designing primers and probes for analyzing  
 PT diseases associated with cytosine methylation state e.g. cancer.  
 XX  
 PS Claim 1; SEQ ID NO 409; 27pp; English.  
 XX  
 XX The invention relates to a nucleic acid comprising a sequence of 18  
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with  
 CC bisulphite, of genes associated with tumour suppression and oncogenes  
 CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and  
 CC 500 are missing from the sequence listing) sequences (S8) and sequences  
 CC complementary to (S8). The nucleic acid may be a peptide nucleic acid-  
 CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of  
 CC probes for detecting the cytosine methylation state and/or single  
 CC nucleotide polymorphisms and also to be used in an array for analysing  
 CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The  
 CC probes can also be used in a method for ascertaining genetic and/or  
 CC epigenetic parameters for the diagnosis and/or therapy of existing  
 CC diseases or the predisposition to specific diseases, by analysing  
 CC cytosine methylations. The parameters may be compared to another set of  
 CC genetic and/or epigenetic parameters, the differences serving as basis  
 CC for diagnosis and/or prognosis events which are disadvantageous to  
 CC patients. The present sequence is one of the 533 genomic sequences  
 CC derived from tumour suppressor genes and oncogenes. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 SQ Sequence 7442 BP; 1655 A; 214 C; 1934 G; 3636 T; 0 U; 3 Other;

Query Match 0.7%; Score 67.6; DB 4; Length 7442;  
 Best Local Similarity 45.4%; Pred. No. 0.0079;  
 Matches 328; Conservative 0; Mismatches 384; Indels 10; Gaps 2;

QY 9012 TGTGATGTAATATGTCATTTAAGAAAAATGTAGAGCAAAATCGAGAGCATGCTAT 9071  
 DB 2679 TATTAAAAAATAAATACTAATAAAAAATAAATACTAATAAAAAATGAAATTT 2620  
 QY 9072 TGTAGAACTACTCGAA-----AACAGATATCAAGCATTTTCAGAGCAAAAGTAATA 9127  
 DB 2619 AAAAAAATAAATACTAATAAAAAAATAAATACTAATAAAAAAATAAATTC 2560  
 QY 9128 TTCTTGAAAAAGAGAGCGCTGACGCTGACATATGATGATACATTTCCAAATG 9187  
 DB 2559 TAAAAAATAAATACTAATAAAAAAATAAATACTAATAAAAAAATAAATTC 2500  
 QY 9188 AGATGATATTTAAAAATTGGCAAGAGATGATGCTTTCTCAATTAATACCAAAAT 9247

DB 2499 TAAAAATAAAAATAAATTCGAAAAAATAAATAAATAAATAAATAAATC 2440  
 QY 9248 CAAAAAATAATATTAATCTTAGCATCAAGTAGTGAATGTAATGTTGATGGGTGG 9307  
 DB 2439 TAAAAAATAAATACTAATAAAAAAATAAATACTAATAAAAAAATAAATCTAA 2380  
 QY 9308 CTGAAGCAAGAGTGTCAGAGCCAAAGCGACAGTTAGTGAAGATCAAAATAATAGAA 9367  
 DB 2379 AAAAAAATAAATCAATATTAATAAAAAAATAAATAAATAAATAAATAATCTAA 2320  
 QY 9368 CTAAATATGTTGATTTAGCAGGAAAAATTAATAACAGAGGAAACATCAATATATGCCG 9427  
 DB 2319 AAAAAAATAAATACTAATAAAAAAATAAATAAATAAATAAATAAATAAATCTAA 2260  
 QY 9428 GATATGATAAAAATTATATATATAGTAAAGCAAAATCTAAGGCTATTGCGATGCCAAA 9487  
 DB 2259 AAAAAATTAACCTCTAATAAAAAAATAAATAAATAAATAAATAAATAAATCTAATA 2200  
 QY 9488 GTCATGCTGACGCTGCTTGGCAACTGCCACTATTGAAAAAATGAAGTAAATTT---- 9543  
 DB 2199 AAAAAAATAAATCAATCTAATAAAAAAATAAATAAATAAATAAATAAATAAATAA 2140  
 QY 9544 --AATATGCAATCCGAAATTTAAATAATCTGGCAAGTTGGAAAGGAAAGCTAATA 9601  
 DB 2139 AAAAAAATAAATACTAATAAAAAAATAAATAAATAAATAAATAAATAAATCTAAT 2080  
 QY 9602 AAAAAAGCTCGATGAGATCTAATCAGGTAGACTGATATACGATAAATATACATGCAAT 9661  
 DB 2079 AAAAAATTAATAATCTAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 2020  
 QY 9662 CTCTGAAAAAGCATACAAAAAATTGACATATCAATCAAGAGAGAGAAAGGAAA 9721  
 DB 2019 AATCTAAAAAACACTAATAATAAATAAAGCACTCTAATAAAAAAATAAATAAATAA 1960  
 QY 9722 AA 9723  
 DB 1959 AA 1958

RESULT 36  
 AAL53680/C  
 ID AAL53680 standard; DNA; 15507 BP.  
 AC  
 XX AAL53680;  
 AC  
 DT 07-FEB-2003 (first entry)  
 XX  
 DB Genomic DNA sequence part 1 of the human ch10-ARPR protein.  
 XX  
 KM Cytostatic; chromosome 10 arginine-rich protein; ch10-ARPR; ARPR; cancer;  
 KM gene therapy; human; gene; de.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT 5'UTR  
 FT  
 FT CDS  
 FT  
 FT exon  
 FT  
 FT intron  
 FT  
 FT exon  
 FT  
 FT intron

Location/Qualifiers  
 927..962  
 /tag= a  
 963..13435  
 /tag= b  
 /product= "Part of the ch10-ARPR human protein"  
 /note= "No stop codon. Coding sequence contains 2  
 introns"  
 963..1077  
 /tag= c  
 /number= 1  
 1078..10639  
 /tag= d  
 /number= 1  
 10640..10767  
 /tag= c  
 /number= 2  
 10768..13294

```

FT      /tag= d
FT      /number= 2
FT      exon      13295..13435
FT      /tag= c
FT      /number= 3
PN      WO200279246-A2.
PD      10-OCT-2002.
PF      26-MAR-2002; 2002WO-EP003395.
PR      30-MAR-2001; 2001US-0280673P.
PR      24-MAY-2001; 2001US-0293453P.
PA      (GENE-) GENEPROT INC.
PI      Bouguetelret L, Niknejad A, Bairoch A;
DR      WPI; 2003-040654/03.
XX      New isolated chromosome 10 arginine-rich protein related polypeptides,
XX      useful for detecting and/or monitoring and creating conditions involving
XX      aberrant expression of ARPR or uncontrolled growth of tissues, such as
XX      cancer.
XX      Claim 8; Page 55-65; 82pp; English.
XX      The invention relates to isolated chromosome 10 arginine-rich protein
XX      related polypeptides (ch10-ARPR), or their fragments. The polypeptides,
XX      nucleic acids and antibodies are useful for detecting and/or monitoring
XX      and creating conditions involving aberrant expression of ARPR or
XX      uncontrolled growth of tissues, such as cancer. The polypeptides are
XX      useful as hybridisation probes, in chromosome and gene mapping, for the
XX      generation of antisense RNA or DNA and in tissue or cell typing. The
XX      methods are useful for detecting and measuring quantities of ARPR in
XX      tissues and biological fluids. The host cells are useful for replicating
XX      ARPR transcripts or expressing the ARPR proteins or polypeptides. The
XX      polynucleotides of the invention can be used to treat disorders by gene
XX      therapy. This sequence represents the genomic DNA sequence, part 1, of
XX      the human ch10-ARPR protein of the invention
XX      Sequence 15507 BP; 4046 A; 3576 C; 3266 G; 4618 T; 0 U; 1 Other;
SQ
Query Match      0.7%; Score 66.2; DB 8; Length 15507;
Best Local Similarity 50.8%; Pred. No. 0.018;
Matches 158; Conservative 0; Mismatches 153; Indels 0; Gaps 0;
QY      7105 GTAGGAGAAAAAGTAAATGTAAACGGCAAAAGAAAAAATTTTAAAGCAACAGCAGCA 7164
DB      12860 GAAAGAAAGAAAGAAATGATGAAAGAAAGAAAGAAAGAAAGCAATGCAGCAGCAGCA 12801
QY      7165 AATGCGAGATCGAGAGCAGCAATCGAGCCAAATGCTGTGTAATTAATTTGGAACA 7224
DB      12800 GCAGCGCGAGCAAGACAGAGAAAGAAAGAAAGAAAGAAAGAAAGCAAGCA 12741
QY      7225 GCTGTGAAGATGAAAGAAATTTGTAAGAAAGAAAGAAAGAAAGTTTAAACTTTAGAC 7284
DB      12740 GCAGCGCGAGCAAGCAAGCAGCAAGAAAGAAAGAAAGAAAGCAAGCAAGCA 12681
QY      7285 GAAAGTTAAAGAAAGCAAGTAAAGTAAATGATCTTACGAAAAAACTTTACATCA 7344
DB      12680 GCAGCGCGAGCAAGCAATTAAGACGATTAAGAAAGCAAGCAAGCAAGCAAGCA 12621
QY      7345 GCAGGATTTTCTACAGAGATCTTGTAAAGCCGATAGAGAGATCTCAGGAGAA 7404
DB      12620 GCAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 12561
QY      7405 GGAATTTAAAGC 7415
DB      12560 GCAGCGAAGC 12550

```

```

RESULT 37
ADQ62833
ID      ADQ62833 standard; RNA; 1000 BP.
XX      AC      ADQ62833;
XX      DT      07-OCT-2004 (first entry)
XX      DE      Homopoly-A contaminant for RNaseH activity assay.
XX      KW      ss; nuclease-mediated cleavage; target nucleic acid; RNaseH activity;
XX      KW      fluorophore; fluorescence quencher; fluorescent signal; HIV;
XX      KW      reverse transcriptase.
XX      OS      Synthetic.
XX      PN      WO2004059012-A1.
XX      PD      15-JUL-2004.
XX      PF      22-DEC-2003; 2003WO-US040879.
XX      PR      23-DEC-2002; 2002US-0436125P.
XX      PA      (AMHP ) WYETH.
XX      PI      Olson MW, O'Connell JF;
XX      DR      WPI; 2004-543471/52.
XX      PT      Detecting a nuclease-mediated cleavage of a target nucleic acid, useful
XX      PT      for detecting and monitoring RNase H activity, comprises hybridizing a
XX      PT      target nucleic acid to a fluorescently labeled oligonucleotide probe.
XX      PS      Example 1; SEQ ID NO 4; 61pp; English.
XX      The invention relates to a method of detecting a nuclease-mediated
XX      CC      cleavage of a target nucleic acid or measuring a RNaseH activity of an
XX      CC      agent by hybridizing a target nucleic acid to a fluorescently labelled
XX      CC      oligonucleotide probe complementary to the target nucleic acid and
XX      CC      containing a fluorophore at one terminus and a quenching group at the
XX      CC      other terminus and contacting the probe-target hybrid with an agent
XX      CC      having nuclease activity. When the oligonucleotide probe is unhybridised
XX      CC      to the target nucleic acid, the probe adopts a conformation that places
XX      CC      the fluorophore and quencher in such proximity that the quencher quenches
XX      CC      the fluorescent signal of the fluorophore. Formation of the probe-target
XX      CC      hybrid causes sufficient separation of the fluorophore and quencher to
XX      CC      reduce quenching of the fluorescent signal of the fluorophore followed by
XX      CC      contacting the probe-target hybrid with an agent having nuclease activity
XX      CC      to selectively cleave the target nucleic acid and thus release the intact
XX      CC      probe. Detection of the release of the probe is by measuring a decrease
XX      CC      in the fluorescent signal of the fluorophore as compared to the signal of
XX      CC      the probe-target hybrid. The methods are useful for detecting and
XX      CC      monitoring RNaseH activity in real time. The methods are also useful for
XX      CC      monitoring enzymatic degradation of an RNA-DNA duplex. In an example of
XX      CC      the invention, the activity of the RNaseH from E. coli or HIV reverse
XX      CC      transcriptase is measure by the method of the invention. In order to test
XX      CC      the specificity of the RNaseH cleavage of the target-probe complex, a DNA
XX      CC      or RNA contaminant is added to the assay reaction. This sequence
XX      CC      corresponds to the single stranded homopolymeric polyA RNA contaminant
XX      CC      for testing this activity.
XX      SQ      Sequence 1000 BP; 1000 A; 0 C; 0 G; 0 T; 0 U; 0 Other;
Query Match      0.7%; Score 65.8; DB 12; Length 1000;
Best Local Similarity 41.8%; Pred. No. 0.0098;
Matches 403; Conservative 0; Mismatches 562; Indels 0; Gaps 0;
QY      8759 AAATGAGATTAATTGAAGCATTGCAAGATATATGATTTGATGAGTCAACGAG 8818
DB      1 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 60
QY      8819 ACGGAAACAAAGAGGTCTTGTGCGAAGAAAGATGTTCTGTGAAAAAATACATTTCAG 8878

```



```

Db      820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
Qy      8999 TTGGATTGGAAATGTTGATGTAATTAATGTGATTAAGAAAAATGTAGAGCCAAATCG 9058
Db      760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
Qy      9059 GAAGACATGCTATTGTAGAACTACTGGAACAAAGAAATATCAGCATTTACAGAGCA 9118
Db      700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641
Qy      9119 AAGTAATATTTCTTGAAAAAGAGACGCTGACAGCTGACAGTCAATATCGAATGTACCA 9178
Db      640 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 581
Qy      9179 TTTCCATGAGATGATATTAAAAATTTGGCAAGCATGATCTTCTCATTAATTA 9238
Db      580 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 521
Qy      9239 CCAAAATTCAAAAAATATATTACTTTAGCATCAAGTAGATGAAATCGAATGTTTC 9298
Db      520 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 461
Qy      9299 ATGGGTGGCTGACCAAGAGGTGACGAGCCAAAGCGACAGTGTAGTAAGAAATCAA 9358
Db      460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
Qy      9359 TAAATAGAACTAATAATGTTGATTAGCAGAAAAATTAACAGAGGAAACATCATG 9418
Db      400 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341
Qy      9419 TATATCCGATATGATTAATAATTAATATAGTAAGCAAAATTTAGAGCTATTTGCGG 9478
Db      340 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 281
Qy      9479 ATGCCAAAGTCATGCTGACGCTGCTTCCGCAATCTCCACTATTGCAAAAAATGAATGA 9538
Db      280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221
Qy      9539 AATTATATGATGATCCAGAAATTTAAATATCTGCAAGATTGAGAGGAAAGCTA 9598
Db      220 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 161
Qy      9599 ATAAAAAAACGTCGGTAGATCTAATCAGGTAGACTGTATACGATTAATATACGTGC 9658
Db      160 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101
Qy      9659 ATTCTTCTGAAAAAGATACAAATAATGACATATCAATCAAGAGAGAAAAAGGA 9718
Db      100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 41
Qy      9719 AAAAA 9723
Db      40 AAAAA 36

RESULT 39
AAA70105
ID      AAA70105 standard; DNA; 5940 BP.
XX      AAA70105;
AC      07-NOV-2000 (first entry)
DE      Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:238.
XX      Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
KW      antimalarial; malaria; protozoicide; infection; insecticide; ds.
XX      Plasmodium falciparum.
OS      Plasmodium falciparum.
XX      Plasmodium falciparum.
PN      W0200025728-A2.
XX      11-MAY-2000.
PD

```

```

XX      05-NOV-1999; 99WO-US026796.
PF      05-NOV-1998; 98US-0107131P.
XX      05-NOV-1998; 98US-0107131P.
XX      (HOFF/) HOFFMAN S.
XX      (CARU/) CARUCCI D.
XX      (GARD/) GARDNER M.
XX      (VENT/) VENTER J C.
XX      Hoffman S, Carucci D, Gardner M, Venter JC;
PI      WPI; 2000-365347/31.
DR      Proteins encoded by chromosome 2 of the human malarial parasite,
XX      Plasmodium falciparum, useful as antimalarial vaccines and in the
PT      diagnosis of P.falciparum infection.
XX      PS      Disclosure; Page 460-462; 577pp; English.
XX      CC      The present invention describes proteins and their fragments (I) encoded
CC      by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC      Also described are: (1) nucleotide sequences (II) encoding (I); and (2)
CC      vaccines against P. falciparum infection comprising (I) or (II). (I) and
CC      (II) are useful for the development of vaccines against P. falciparum
CC      infection. (I) and polyclonal antisera or a monoclonal antibody raised to
CC      immunogens comprising the sequences of (I), are useful in the detection
CC      of infection with P. falciparum. Furthermore, (I) (especially when they
CC      are refined or secreted or membrane proteins) can aid the identification
CC      of drugs to treat or prevent P. falciparum infection, or they can be used
CC      to identify drug resistance in P. falciparum. Sequencing of the
CC      Plasmodium chromosome 2 and the subsequent identification of proteins
CC      encoded by it will help to expand our understanding of parasite biology,
CC      a process hampered by the complexity of the parasitic life cycle, and
CC      provide new targets for vaccine and drug development. Parasite resistance
CC      to drugs and mosquito resistance to insecticides have led to a resurgence
CC      of malaria in many parts of the world, and there is a pressing need for
CC      vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
CC      represent nucleotide and protein sequences given in the present
CC      invention, but which are not specifically mentioned within the
XX      specification
XX      SQ      Sequence 5940 BP; 3106 A; 343 C; 879 G; 1612 T; 0 U; 0 Other;
XX      Query Match 0.7%; Score 65.4; DB 3; Length 5940;
XX      Best Local Similarity 44.7%; Pred. No. 0.02;
XX      Matches 472; Conservative 0; Mismatches 571; Indels 14; Gaps 5;
Qy      10 ATCAAAATTAAGCTTCAGAGGACCAAGAGGATATCATGATTTCAAAAAAGTTTAAAG 69
Db      136 ATAGAAACTTAAGAGAAAAAGTAAATTAATAATTAATTAATCTTATGATGATTAAT 195
Qy      70 ATTTGGAGTTGTGATTAACTATGACGCGGTGAGGCTTAATGATCAATCACCGGACT 129
Db      196 ATTTATGAGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 255
Qy      130 GAGAAATTTGGAACAAATAAGAAAAAGATAATGTTTATGACATTACTCAACAAG 189
Db      256 AATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 315
Qy      190 ATTCAGGGGAGAACGCTTTTAACGTTTAAATAGATTGCTTTAAGAAAAATAATATA 249
Db      316 AATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 374
Qy      250 GCAATATCTATTTTGGGAAAAAGATAGTAGCGGGGTAAATATCTTTTAACTTTGTC 309
Db      375 TAAAGCAATGAGCTGAGAAACACAGCTTAAGATACATTAATTAAGTCATTAGTTCGTC 434
Qy      310 AATGAAAAAATGAAGTAGATGATTAATCAACGAATTCGAGAAATTAATTAATTAATGA 369
Db      435 GATTAATTAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 494
Qy      370 AATTATATTTTCTTAAGCTCGAAGGAGATGCAATGAGAAAAAATGCAATGCT 429

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2005, 03:43:28 ; Search time 22519 Seconds  
(without alignments)  
20207.432 Million cell updates/sec

Title: US-10-647-057-8  
Perfect score: 9726  
Sequence: 1 atgagcgcgcacataataa.....gagaaaaaggagaaaaatga 9726

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hlc:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gse1:\*  
10: gb\_gse2:\*  
11: gb\_gse3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	104	1.1	1566	10	CG57757 P053-1-D0
C 2	100.8	1.0	1811	10	CG575732 P048-4-G0
C 3	100.8	1.0	1917	10	CG575732 P048-4-G0
C 4	100	1.0	1452	10	CG575732 P048-4-G0
C 5	99.8	1.0	1843	10	CG575732 P048-4-G0
C 6	97.8	1.0	1491	10	CG575732 P048-4-G0
C 7	97.4	1.0	1451	10	CG575732 P048-4-G0
C 8	97.2	1.0	1434	10	CG575732 P048-4-G0
C 9	96.8	1.0	1519	10	CG575732 P048-4-G0
C 10	96.6	1.0	1691	8	DR124185 49179719
C 11	96.2	1.0	2001	10	CG575732 P048-4-G0
C 12	95.4	1.0	1547	10	CG575732 P048-4-G0
C 13	95.2	1.0	1784	10	CG575732 P048-4-G0
C 14	95	1.0	1721	8	DR145459 49056454
C 15	93.8	1.0	1482	10	CG575732 P048-4-G0
C 16	93.2	1.0	1416	10	CG575732 P048-4-G0
C 17	93.2	1.0	1531	10	CG575732 P048-4-G0
C 18	93	1.0	1594	10	CG575732 P048-4-G0
C 19	92.4	1.0	1654	10	CG575732 P048-4-G0
C 20	91.8	0.9	1256	10	CG575732 P048-4-G0
C 21	91.8	0.9	1485	10	CG575732 P048-4-G0
C 22	91.8	0.9	1808	8	DR125630 49056520

C 23	91.6	0.9	1528	10	CG575732 P048-4-G0
C 24	91.4	0.9	1453	10	CG575732 P048-4-G0
C 25	91.4	0.9	1829	10	CG575732 P048-4-G0
C 26	91.2	0.9	1981	10	CG575732 P048-4-G0
C 27	90.6	0.9	1507	10	CG575732 P048-4-G0
C 28	90.2	0.9	1401	10	CG575732 P048-4-G0
C 29	90.2	0.9	1407	10	CG575732 P048-4-G0
C 30	90.2	0.9	1618	10	CG575732 P048-4-G0
C 31	89.6	0.9	1824	10	CG575732 P048-4-G0
C 32	89.2	0.9	1932	8	DR141720 49300102
C 33	88.6	0.9	1434	9	CG575732 P048-4-G0
C 34	87.8	0.9	2071	8	DR147304 49027161
C 35	87.6	0.9	1457	10	CG575732 P048-4-G0
C 36	87.4	0.9	1428	10	CG575732 P048-4-G0
C 37	87.2	0.9	1139	1	AJ927291 AJ927291
C 38	87.2	0.9	1512	10	CG575732 P048-4-G0
C 39	87.2	0.9	1783	8	DR141924 49190649
C 40	87	0.9	1650	10	CG575732 P048-4-G0
C 41	86.6	0.9	1438	10	CG575732 P048-4-G0
C 42	86.6	0.9	1816	10	CG575732 P048-4-G0
C 43	86.4	0.9	1314	10	CG575732 P048-4-G0
C 44	86.4	0.9	1616	10	CG575732 P048-4-G0
C 45	86.4	0.9	1753	8	DR131208 49217594

#### ALIGNMENTS

RESULT 1  
CG57757/57/c  
LOCUS P053-1-D07.zc Ppa EcORI BAC library Pristionchus pacificus genomic,  
DEFINITION genomic survey sequence.  
ACCESSION CG57757  
VERSION CG57757.1 GI:37986636  
KEYWORDS GSS.  
SOURCE Pristionchus pacificus  
ORGANISM Pristionchus pacificus  
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Nematoda; Pristionchus.  
AUTHORS Srinivasan,J., Stutz,W., Jesse,T., Wieggers-Perepolte,L., Jansen,K.,  
Buntjer,J., van der Meulen,M. and Sommer,R.J.  
An integrated physical and genetic map of the nematode Pristionchus  
pacificus  
TITLE Mol. Genet. Genomics 269 (5), 715-722 (2003)  
JOURNAL 12884007  
PUBMED Contact: Sommer RJ  
COMMENT Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
Class: BAC ends.  
FEATURES  
source  
1..1566  
Location/Qualifiers  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Ppa EcORI BAC library"  
/note="The library was generated by a partial digest of  
the genomic DNA with EcORI and cloning into the BAC  
vector."  
ORIGIN  
Query Match 1.1%; Score 104; DB 10; Length 1566;  
Best Local Similarity 37.0%; Pred. No. 4,9e-10;  
Matches 482; Conservative 0; Mismatches 820; Indels 0; Gaps 0;  
QY 8422 AACGAGATTATGAAATATGATTCACAAAATAATTTGAGAGTCAATGCTCT 8481

[illegible][illegible]

RESULT 2  
 CG753732/c  
 LOCUS  
 DEFINITION  
 CG753732  
 P0484-G03.ya Ppa ECoRI BMC Library  
 genome survey sequence.  
 CG753732  
 CG753732.1 GI:37978509  
 GSS.  
 Pristionchus pacificus  
 Pristionchus pacificus  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
 Neodiplogasteridae; Pristionchus.  
 1 (bases 1 to 1811)  
 Srinivasan,U., Sins,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,  
 Buntjer,J., van der Meulen,M. and Sommer,R.J.  
 An integrated physical and genetic map of the nematode *Pristionchus*  
*pacificus*  
 Mol. Genet. Genomics 269 (5), 715-722 (2003)  
 1811 bp DNA linear GSS 24-OCT-2003  
 JOURNAL  
 PUBMED  
 12884007  
 COMMENT  
 Contact: Sommer RJ

Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel.: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@uebingen.mpg.de  
Class: BAC ends.

```

FEATURES
    source
        location/Qualifiers
            1..1811
                /organism="Pristionchus pacificus"
                /mol_type="genomic DNA"
                /strain="California"
                /db_xref="taxon:54126"
                /clone_lib="Ppa Scori BAC library"
                /note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

```

Query Match	1.04;	Score 100.8;	DB 10;	Length 1811;
-------------	-------	--------------	--------	--------------

Best Local Similarity 34.2%; Pred. No. 2.3e-09;

Matches 567; Conservative 0; Mismatches 1090; Indels 2; Gaps 1/

8068 AAAGTTAGTACAAATCCGAGAGAGTATTTTCATGCAGATAATGTGATATGGAGCA 8127

---

---

Db 1751 AAANNANAAAAAANNAAAANNNAANNNAANNNAANNNAANNNA 1692

8128 ACACATAAATGAAGTACACGAGTTTCTAAGCAGTACACGTTCTGTATTGGAGGA 8187

\_\_\_\_\_

[illegible]

8188 GTTGCAGTTCACCAAGGCAGAGCTACTGCTGCAGCTAAACTATGCTAGAGATTGAGCAA 8247

— 1 —

Ph  
1631 AANTTAAATAAATNTANVAAAATTANAATAATTAATTTAAAAA 1572

[illegible]

0240 GGAAATTGTTTACGAAACAAATCGAATTCTTAAAGTACGAGGCTTTGGAATGAA 0307

[illegible]

1513

[illegible]

OY	8308	GATAAAGTAACTGCTAAATCTTTGTAATATCAGGAAATGAGAGGAAATTGCCGAGCA	8367
Db	1511	AAA	1452
OY	8368	GGAGTGAATACCTCTCAACAGCACAAAAGTATATCTGATCCGTAGTTCGTTACGAAAGCA	8427
Db	1451	ANANANAAA	1392
OY	8428	GATTATGAAAATTAATGATTACACAAAAAATATATTTCAGAGTCATGCTCTTGCTTTA	8487
Db	1391	ANNNAAA	1332
OY	8488	AATGATCAAAAGATGAAGCGAATATGTGAATCTTTAGCGGTAGCCGGTGTGCATGCACA	8547
Db	1331	AAA	1272
OY	8548	GGAACAAACAAGCATTTACAGAGATCAACAAGTTAATCTTCACACTGTAAATGAGGA	8607
Db	1271	AAA	1212
OY	8608	AACGTATCTCAACTTCGTGCAAAAGCTTTGGCTAAAAATGAAAAATTAGAAATGTAAA	8667
Db	1211	AAA - ANAAAAA	1154
OY	8668	GGAATCGAGAGACCTTAGTCGAGCGGAAAACAGACCCGTTGAAATTTATCAAGAGT	8727
Db	1153	AAA	1094
OY	8728	ACTACAGAGCATTTGGTTCAGAGAAATTGGGAAATTGAGATTAATTGAACGATTGCA	8787
Db	1093	ANAAA	1034
OY	8788	AGAGATTAATACATTTGTAAAGAGTCAACGAGACGAAACCAAGAGGTCTTGTGCGAAG	8847
Db	1033	AAA	974
OY	8848	AATGTATTTCTGTGAAAAATATCAATTTCAAGGGAACAAAATCATCTTGAAGATTA	8907
Db	973	AAA	914
OY	8908	GCCAGAAATTTGTGAACCGAAGGTAAATGTAGTGTCTTTGAATGAACTTGATGTGAT	8967
Db	913	AAA	854
OY	8968	CTACAGAGAAAGTGTGTGCTATGTGTGAATTTGTATTTGAAATTTGATGATTAAT	9027
Db	853	AAA	794
OY	9028	GTGATTAAGAAAAATGTAGAAACCAAAATCGAAGACATGCTTATTTAGAAACTATCGGA	9087
Db	793	ANNNNAAA	734
OY	9088	AAACAGAAATTCAGCAATTTCAAGAGCAAAAGTAAATATTTTGGAAAAAGAGCGCT	9147
Db	733	AAA	674
OY	9148	GCAGCTGACGTCAATATGTCAATGTACACATTTCCAATGAGATGATTAATTAATTTG	9207
Db	673	NAAA	614
OY	9208	GCAAGCAGTATGCACTCTTCTCAATTAATTAACAAAAATTCAAAAAATTAATTAATTTA	9267
Db	613	AAA	554
OY	9268	GCATCAATATGTGAATCGAATGTGAATTTCAATGGGTGCTGAAGCAAGGTGCAGGA	9327
Db	553	NAAA	494
OY	9328	GCCAAAGCGAGTATGTATAAGCAATCAAAATTAATAAGCTAAATATGTTGATTTAGCA	9387
Db	493	AAA	434
OY	9388	GGAATAATTTAAACAGAGGAAACATCAATGTATATGATCCGAGATATGATTAATAATTAAT	9447

[illegible]

### RESULT 3

LOCUS	CL081899	1917 bp	DNA	linear	GSS 31-DEC-2003
DEFINITION	CH21c-164124_Sps.1 CH21c <i>Xenopus tropicalis</i> genomic clone				
DESCRIPTION	CH21c-164124, genomic survey sequence.				

## KEYWORDS

SOURCE *Xenopus tropicalis* (western clawed frog)

1. (bases 1 to 1917)  
Acanthopneuste, Acanthopneuste, Acanthopneuste.

### TTT.R

**JOURNAL**  
**Unpublished (2003)**

Insert Length: 175000    Std Error:

High quality sequence stop: 1140.

```
/strain="Nigerian frog"
```

```

/c1one lib="CH216"

```

## ORIGIN

Q10000: Match 1 08: Score 100 8: DB 10: Length 1817:

Best Local Similarity 42.8%; Pred. No. 2.3e-09;

27

DB 327 AAAGANCGAAGGAAAGAAAC 386

OY		8481	TGCTTAAATGTATAAAGGAATTGAAGCGCAATATAGACTTTTAGCGTGACGGGTGC	8540
Db		387	ACAAAAAACAAAACAAAAAAAAAAAAAAAAACAACAACCAACAAAAACAAAAA	446
OY		8541	TGCACAGAACAACAAGCATTTACGAGATCAACCAAGTTAACTTCAACACTGTAA	8600
Db		447	AAAAAACAAACCAAAGAAAGNANNAAGAAAAAGAAAGAAAGAAAAAGAAAAA	506
OY		8601	TGAGAGAAACGTATCTCACTTCGTCAAAAGCTTTGGCTAAAAATGAAAATTATGAA	8660
Db		507	AAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAGAAAAAAGAAAAAAGAAAAA	566
OY		8661	TGTAAAGGAAGCTGAGAGAGCCTTAGTGGAGCCGGAACACAGAGCCGTGAAATTATAC	8720
Db		567	AAAGAA	626
OY		8721	AAAGAGTACTACAGAGCATTTGTTGCAGAAATTGGGAAATTTGAGATTAATTAGAAC	8780
Db		627	AAGAAAAAGAAAAAAGGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	686
OY		8781	GATTGCAAGATTAATACGATTGTAGAGTCAACGAGACCGAACCAAGAGGTCTTGT	8840
Db		687	AAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	746
OY		8841	CGAAGAGAAATGTTATTTCTGTGAAAAATACAAATTCAGGGAACCAAAATCATCATTA	8900
Db		747	AAAAAGAAAAAGGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAGAA	806
OY		8901	AGATTAAGCCAGAAATTTGTTGGACCAGAGTGTAAATAGATGTTGAATGAACCTTA	8960
Db		807	AAAAAACAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGCAAGAAAAAAGAAAA	866
OY		8961	TGTAGATCTACAGAGAAAAAGTGGTGGCTATGTGTGAATTGGTAATTGGAATGTTAGT	9020
Db		867	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGATTA	926
OY		9021	AAATATGTGATTAGAAAAAATGTAGAGCCAAATTCGAGAGCAATGCTATTTATGAAAC	9080
Db		927	AAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA	986
OY		9081	TACTGGAACACAGAGATATCAAGCATTTTACAGAGCAAAAGTAAATTTCTTGGAAAG	9140
Db		987	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA	1046
OY		9141	AGACGCTCAGCTGCAGTGCATATCGAATGTACACATTTCCAATGAGATGATATTAA	9200
Db		1047	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	1106
OY		9201	AAATTTGCCAAGCGATATGCTCTTCAATTAATACCAAAATTCAAAAATATAT	9260
Db		1107	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA	1166
OY		9261	TACTTAGCATCAAGTAGTGAATCGAATGTGAATGTTCATGGCGGTGAGCAAGAG	9320
Db		1167	AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	1221
OY		9321	TGCAGAGACCAAGCGACAGATTAGTAAAGATCAAAATTAATAGACTTAATATGTTGA	9380
Db		1222	GAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA	1281
OY		9381	TTTACAGGAAAAATTTAAACAGAGGAAACATCAATGTATATGCGGATATGATAAAA	9440
Db		1282	GAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA	1341
OY		9441	TTATATAATTAAGTAAACAATTTCTAAGGCTATTGCGAGTGCAAAAGTCAATGCTGCAC	9500
Db		1342	AAGCAAGAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAGAA	1401
OY		9501	TGCTTCGGCACTGCCACTATTTGAAGAAAAATGAATGAATTTAATTAATGCGATCCGAG	9560
Db		1402	AAAAAAGAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAGAAAAAAGAAAA	1461
OY		9561	ATTTAAAAATTAATCTTGGCAGATTTGGAAGGAAAGCACTAATAAAAACGTGCGTAGATC	9620

Df	1462	AGAAAAAAGAAAAAAGAAACCAAAAAAAAAAGAAAAAAGAAAAAAGAAAAG 15211
Oy	9621	TAAATCAGTAGACTGTATACGATTAATAATATACATGGCATTTCTTGAAAAAGCATCAA 96808
Df	1522	AAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAGAAAGAAAAAGAAGAAA 15811
Oy	9661	AAATTGCATTCATTCACAAGAGAGAGAAAAAGCGAAAAAA 9723
Df	1582	AAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1624
 RESULT 4 CL125503		
LOCUS	CL125503	1452 bp DNA linear GSS 05-JAN-2004
DEFINITION	ISB1-87B20.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-87B20,	
ACCESSION	CL125503	genomic survey sequence.
VERSION	CL125503.1	GI:40619138
KEYWORDS	GSS.	
SOURCE	Xenopus tropicalis (western clawed frog)	
ORGANISM	Xenopus tropicalis	
REFERENCE	Xenopus tropicalis Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana. 1 (bases 1 to 1452) Kremetzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R. A physical map of the xenopus tropicalis genome Unpublished (2003) Contact: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@wustl.edu Insert Length: 75000 Std Error: 0.00 Seq primer: T7 TAATACGACTCATATATAGG Class: BAC ends High quality sequence start: 672 High quality sequence stop: 905. Location/Qualifiers 1..1452 /organism="Xenopus tropicalis" /mol_type="genomic DNA" /db_xref="taxon:9364" /clone="ISB1-87B20" /clone_1fb="ISB1" /note="Vector: pBelOBAC11, ISB-1 Xenopus tropicalis BAC Library Segment 1"	
FEATURES source		
1..1452		
Query Match 1.0%; Score 100; DB 10; Length 1452;		
Best Local Similarity 41.3%; Pred.No.3,1e-09;		
Matches 558; Conservative 0; Mismatches 792; Indels 1; Gaps 1;		
Oy	8374	AATACTTACAGCAAAAGTAAATCTGAATCCGTAGTTCGTTACGAAGCAAGATTAT 84333
Df	23	AAAAAAAAACCACAAAAAAGAAAAAAGAAAAAAGAAAAAAGCAACAAAAAGAACAAACA 82
Oy	8434	GAAATATATGTTACACAAAAAATATATTTCAGAGTCATGCGCTTCGTTTAAATGAT 84933
Df	83	AAAAAAAAAGATTAATAAAAAAAAAANGACACACAAATATAAGAGAGAAAAAANAANA 142
Oy	8494	ACAAAGATGAAGCAATATAGAAATCTTTACGCGTAGCCGCTGTGCATGCAACAAGACA 85533
Df	143	AAAAAAAAAAAAAAAAAAAAAACCCTTAATAAGANNNNNNNNNNNNNNAANAAGAAAA 202
Oy	8554	AACAAAGCATTTACGAGATCAACAAGTTAACTTACACTGTAAATGAGAAACGTA 86133
Df	203	AAAAAAAAANNAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAANAANAATCT 262
Oy	8614	TCTCACTCTGCGAAAAAGCTTTGGCTTAAATGAAATTTATGGAATGTAAAAAGAACT 86733
Df	263	CTAAAGAAAGAAAGANNNNNNNANNNAANNAANNAANNAANNAANNAANNAANNAAGAAA 322

```

QY 8674 GAGAGAGCTTACTCGAGCGGAACAGACGCCGTTGAAATTTATACAAAGACTACTACA 8733
Db 323 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 382
QY 8734 GGAGCATTTGGTTTCAGGAAATTTGGAAATTTGAGATTAATTAGAAACGATTGCAAGAT 8793
Db 383 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 442
QY 8794 AATACATTGTAAAGCTCAACGAGACGGAACCAAGAGAGTCTTGTGGAAGAAATGTT 8853
Db 443 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 502
QY 8854 ATTTCGTGTAATCAATTTTCAGGGGAAACAATTCATCCATTGAAGTAAAGCCAGA 8913
Db 503 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 562
QY 8914 ATTGTTGAAACCGAAGTGAATGTAGATCTTTGAATGAACTTGATGATGATCTACAA 8973
Db 563 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 622
QY 8974 GGAAGAGTGGTGGCTATGTTGGAATTTGTAATGTAATGTAATTTGTAATTT 9033
Db 623 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 682
QY 9034 AAGAAAAATGTAGAACCCAAATTCGGAAGCATGCTATTGTAGAACTACTGGAACAA 9093
Db 683 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 742
QY 9094 GAATATCAAGCATTTCAAGAGCAAAAGTAAATATTCTTGGAAGAGAGACGCTGACGT 9153
Db 743 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 802
QY 9154 GCAGCTGCATATCGAATGTACATTTCCATGAGATGATTTAAATTTTGCAAAAG 9213
Db 803 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 862
QY 9214 CAGTATGCTCTTCTCAATTAATACCAAAATTCAAAAATTAATTACTTTAGCATCA 9273
Db 863 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 922
QY 9274 AGTAGGAAATCGAATGTGAATGTTTCATGGGGTGGCTGAAGCAAGAGTGCAGAGCAAA 9333
Db 923 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 982
QY 9334 GCGACAGTATAGTAAAGATCAAAATTAATGAATTAATGTTGATTTAGCAGAGAAA 9393
Db 983 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1042
QY 9394 ATTAAACAGAGGAAACATCAATGTAATATGCCGATATGATTAATTAATTAATGAT 9453
Db 1043 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1102
QY 9454 AAGACAAATTTCTAAGGCTATTCGGAATGCAAAAGTCATGCTGAGCTGTTGGCACT 9513
Db 1103 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1162
QY 9514 GCCACTATTGAAAAATGAAGTAAATTTAATATCGATCCGAAATTTTAAAAATAT 9573
Db 1163 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1222
QY 9574 CTGGCAAGATTGAGAGGAAAGCT-AATTAATAAAAGCTCGTGTGATCTAATGAGTGA 9632
Db 1223 AAGAGAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAA 1282
QY 9633 CTGGTAAAGGAAATTAATATCATGCAATCTTCTGAAAAAGCATCAAAAAATTTGACAT 9692
Db 1283 AAGAGAAAGGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAA 1342
QY 9693 TCAATCAAGAGAGGAAAAAGGAAAAAA 9723
Db 1343 AAGAACAAACAAAAAGAAACACAGAAAA 1373

```

```

RESULT 5
AG435185/c 1843 bp DNA linear GSS 21-DEC-2004
LOCUS Mus musculus molossinus DNA, clone:MSMg01-313C22.T7, genomic survey
DEFINITION
ACCESSION AG435185
VERSION AG435185.1 GI:48078248
KEYWORDS GSS.
SOURCE Mus musculus molossinus (Japanese wild mouse)
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Abe, K., Noguchi, H., Tagawa, K., Yuzurika, N., Toyoda, A., Kojima, T.,
Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriwaki, K. and
Shiotohshi, T.
Contribution of Asian mouse subspecies Mus musculus molossinus to
genomic constitution of strain C57BL/6J, as defined by BAC-end
sequence-SNP analysis
Genome Res. 14 (12), 2439-2447 (2004)
JOURNAL PUBMED 15574823
AUTHORS Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE Direct Submission
Submitted (17-NOV-2003) Maehira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenho-chou, Tsukumi-ku, Yokohama, Kanagawa, 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kunita Abe (abe@rc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyada, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBAC3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES
source
1.1843
Location/Qualifiers
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-313C22.T7"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_id="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match 1.0%; Score 99.8; DB 10; Length 1843;
Best Local Similarity 31.5%; Pred. No. 3.6e-09;
Matches 469; Conservative 0; Mismatches 1017; Indels 1; Gaps 1;
QY 8237 AAGTTGAGAGAGAAATTTGTTCAAGAAACAAATCGATTGAATGCAATTTCTAAAGTGAAG 8296
Db 1825 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1766
QY 8297 GTTGGATGAAGTAAAGTACGTCTAATTTCTCTAGTATCAGGAAATGAGAGAGAA 8356
Db 1765 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1706
QY 8357 TTGCGGAGACAGAGTAACTTCTACAGCAAAAGTAATCTGAATCCGTAGTTCGTT 8416
Db 1705 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1646
QY 8417 TAGAAAGAGAGATTATGAAATTAATGATTACACAAAAAATATATTTTCAGAAAGTCAATG 8476

```

Db	1645	NANANANNNANNAAAAAAAAAANNANANANANAAAAAAAAANANANNNNNANNA	1586
Qy	8477	CTCTGCTTTAAATGATACAAGATGAAGCAATATGAATCTTTAGCGTAGCCGGT	8536
Db	1585	NANANANANANANANANANANNNANNNANANANANANANANANANANANNA	1526
Qy	8537	TGCATGCAAGAAACAAGACTTTACGAGTCAACAGTTAACTTCTCACTG	8596
Db	1525	NANNNAN	1466
Qy	8597	TAAATGAGAGAAACGTATCTCAACTTCGTGCAAAAGCTTGGCTAAATGAAATATG	8656
Db	1465	NANANANNNANNAAAAAAAAAANANANANANANANANANANANANANANNN	1406
Qy	8657	GAAATGTAAAAGAACTGAGAGAGCTTAGTCCGAGCCGAAACAGCAGCCGTTGAAAT	8716
Db	1405	ANNNANNA	1346
Qy	8717	ATACAAAGATCTACAGAGCACTGGTTCCAGAAATTGGGAAATTGGAGATAATTAG	8776
Db	1345	NNAANA	1286
Qy	8777	AAACGATTGCAAGAGATAATACGATTGTAAAGTCAACGAGACCGAACCAAGAGAGTC	8836
Db	1285	AAANN	1226
Qy	8837	TTGTCGGAAGAAATGTAATTTCTGGAANAAATACATTTCAAGGGGAAACAATATCATCA	8896
Db	1225	NANNNANNA	1166
Qy	8897	TTGAAGATTAAGCCAGAAATGTTGGAACCGAAGTGAATGTAGATCTTTGAATGAC	8956
Db	1165	AAAAAN	1106
Qy	8957	TTGATGTAGATCTACAGGAAAGATGGTGCTACGTGGAATTGTATTGGAAATGTG	9016
Db	1105	NANNA	1047
Qy	9017	ATGTAAATPATGTGATTAAGAAANATGTAGAAACCAAAATCGAAGACATGCTATTGTAG	9076
Db	1046	ANNNANNA	987
Qy	9077	AAACTACTGGAACAACAAGATATACAGACTTTACAGAGCAAAAGTAATATTCTTGAA	9136
Db	986	AAAAANNA	927
Qy	9137	AAAGGAGCGCTCAGCTGCACTGCATATGCAATGTACACTTTCCATGATGATGATA	9186
Db	926	NANNA	867
Qy	9197	TTAAANATTTGGCAAGCGATGTGCATCTTCTCAATATATACCAAAANATTCAAAAATA	9256
Db	866	AAAAANNA	807
Qy	9257	ATATTACTTAGCATCAAGTAGTGAATGCAATGTGATGTTCCATGGGTGGCTGAGCAA	9316
Db	806	AAAAANNA	747
Qy	9317	GAGGTGCAAGACCAAGCGACGTTAGTGAAGATCAATTAATTAATTAATTAATG	9376
Db	746	AAAAANNA	687
Qy	9377	TTGATTTTGCAGGAAATTTAAACAGAGGGAACAATCAAGTATATGTCGGATTTGATA	9436
Db	686	AAAAANNA	627
Qy	9437	AAATTTATATATTAAGTAAAGCAATTTCTAAGCTATTGGCGATGCCAAAGTCACTGCTG	9496
Db	626	AAAAANNA	567
Qy	9497	CAGCTGCTTCGCAACTGCACTATTTGAAAAAATGAAGTAATTTATATGCGATCC	9556

[illegible]

[illegible]

QY	9694	CAATCAAGAGAGAGCAAAAAA	9723
DB	114	AAAAAAAAAAAAAAAAAAAAAAAAA	85
RESULT 7			
AG311072			
LOCUS			
DEFINITION	AG311072	1615 bp	DNA
ACCESSION	AG311072		linear
VERSION	AG311072.1		GS5 18-DEC-2004
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			
JOURNAL			
PUBMED			
AUTHORS			
REFERENCE			
TITLE			

Db	256	AAAAAAAAAGAAAAAGAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAAGAAAAA	315
Qy	8589	TACAACTGTAATGGAAGAAACGTATCTCAACTTCGTCAAAAGCTTTGGCTAAATATGA	8648
Db	316	GAATAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA	375
Qy	8649	AAATATATGAAATGTAAAAAGAACTGGAGAGCCTTAGTCGAGCGGAAACAGCAGCCGT	8708
Db	376	AAAAAAAAAAAAAAAAATTAATAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA	435
Qy	8709	TGAAATATATACAAAGAGTACTACAGAGCATTTGGTTCAGAGAAATTGGAAATTGAGA	8768
Db	436	TAAAAAATAAAAAAAAAAAAAATTAATAAAAAAGAAAAAAGAAAAATTAAGAGAAAAAA	495
Qy	8769	TAAATTAAGAAACGATTGCAGAGATTAATTCGATTTGAAGTCAACGAGACGAAACCA	8828
Db	496	GAAATTAAGAAAAAGAAAAAAATTAATTAATTAATAAAAAAGAGAGAGAGAGAAAAAA	555
Qy	8829	AGAGAGTCTTGTCCGAGAAAGATGTATTTCTGTGAAAAATCAATTTCAAGGGAGACAA	8888
Db	556	AATTAAGAGAGAGAAAAAGAAAAAGAAAAAGAGAGATTAATAAAAAAGAAAAAGAA	615
Qy	8889	ATCATCATTTGAGATTAAGCCAGAAATGTTGAAACCGAGGTAAATGTAGATCTTT	8948
Db	616	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA	675
Qy	8949	GAAATGAATGTATGTATGATCTACAGAGAAAAAGTGTGCTATGTTGGAATTGTAATTG	9008
Db	676	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATTAATAAAAAAGAAAAAAGAA	735
Qy	9009	AAATGTGATGTAATTAATGTGATTTAAGAAAAATGTAG-AAAGCCAAATTCGAGACATG	9067
Db	736	AAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAATTAAGAAAAAATTAATAAAAAAGAA	795
Qy	9068	CTATGTGTAATCTACTGAAAAACAAGAAATTAACAACATTTACAGACCAAAAGTAATA	9127
Db	796	AAAAATTAATTAATAAAAAAAAAAGAAATGAAAAAAGAAAAAAGAAAAAAGAAAAAG	855
Qy	9128	TTCTTGGAAAAAGAGACGCTGACGCTGACGTGCATATTCGATGTACACATTTCCATG	9187
Db	856	AA	915
Qy	9188	AGATGATTAATTAATAATTTGGCAAGCATATGATCTTCTCAATTAATTAACAAAAAT	9247
Db	916	GAATAAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	975
Qy	9248	CAAAATTAATTTACTTTAGCATCAAGTATGATGCAATGTAATGTCAGTGGCGTG	9307
Db	976	AAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAAAGAAAAAATTAAGAAAAAATTA	1035
Qy	9308	CTGAAGCAGAGTCCAGAGCCAAAGCGACAGTTAGTGTAAAGATCAATTAATTAAGAA	9367
Db	1036	ATTAATAAAAAAAAAAAAAAAAAAGAAAAAGAAAAAAGAAAAAATTAATAAAAAAGAA	1099
Qy	9368	CTAATTAATTTGATTAGCAGAAAAATTTAAACAGAGGAAACATGTAATGTGCG	9427
Db	1096	AAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAATTAAGAAAAAAGAAAAA	1155
Qy	9428	GATATGATTAATAATTAATATATTAAGTAAAGACAAATTTCAAGCTATTCGG---ATGCCA	9488
Db	1156	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAATTAATAAAAAAGAAAAA	1215
Qy	9485	AAAGTCATGCTCAGCTGCTTGGCACTGCCATTAATGAAAAAATGAAGTAATTTTA	9544
Db	1216	AAAGAA	1275
Qy	9545	ATTAATGCGATCGAGATTTAAATAATCTGCAAGATTTGGAAGGAAAGCTATTA	9604
Db	1276	AAATTAATAAAAAAAAAAGAAATGAGAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1335
Qy	9605	AAACGTGCTAGATCTAATCAGGTAGACTGTATACGATTAATATATCATGCAATTTCT	9664

[illegible]

OY	8601	GGAGGAAACGTAATCTCACTTCGTCGCAAAACCTTGGCTAAATATGAAAATTATGGAAA	8660
DB	1195	NAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	1136
OY	8661	TGTAAAGAAAGAACTGGAGAGCCCTTAGTCGAGCGGAAACAGACGCCGTTGAAATTTATAC	8720
DB	1135	ANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	1076
OY	8721	AAAGAGTACTACAGAGCATTGGTTGCAGAAATTGGGAAATTGAGATTAATTAGAAAC	8780
DB	1075	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	1016
OY	8781	CATTGCAAGAGATTAATACGATTGTATAGACTCAACGAGACGGAACCAAGAGGCTCTTGT	8840
DB	1015	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	956
OY	8841	CGAAGAGATGTGTAATTTCTGTGAATAATACAAATTCAGGGGAAACAAATCATCATTTGA	8900
DB	955	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	896
OY	8901	AGATTAAGCCAGAAATTGTTGAAACCGAAGTGTAAATGTAGACTCTTGTGATGAACTTGA	8960
DB	895	ANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	836
OY	8961	TGTAGATCTACAGGAAAAAGTGGCTGCTATGTGTGAAATTGCTATTTGGAAATGTTGATGT	9020
DB	835	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	776
OY	9021	AAATTAATGTGATTAAGAAAAATGTAGAACCAAAATCGAAGACATGCTATTGTAGAAC	9080
DB	775	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	716
OY	9081	TACTGGAAGACAAGATATCAAGCTTTACAAAGCAAAAGTAATATATCTTGGAAAAAG	9140
DB	715	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	656
OY	9141	AGAGCTGACGCTGCAGCTGCAATATCGAATGTACACATTTCCAATGAGATGATTTAA	9200
DB	655	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	596
OY	9201	AAATTTGGCAAGCAGTATGCACTTCTCAATTATTAACCAAAATTCAAAAATTAATAT	9260
DB	595	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	536
OY	9261	TACTTTAGCATCAAGTAGTAGAATCGAATGTGAATGTTTCATGGGGTGCTGAAGCAGG	9320
DB	535	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	476
OY	9321	TGCAGAGCCAAAGCGACAGTTAGTGTAAAGAAATCAATATAATATAGAACTATATATGTTGA	9380
DB	475	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	416
OY	9381	TTTAGCGAGAAAAATTAAACAGAGGGAACATCAATGTATATCCGATATGTATAAA	9440
DB	415	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	356
OY	9441	TTATATATTAAGTAGACAATTTCTAAGGCTATTTGCGGATCCAAAATGTCATGCTGCAGC	9500
DB	355	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	296
OY	9501	TGCTTCGGCAACCTGCCACTATTGAAAAAATGAAGTAATTTATATATGCGATCCGAGA	9560
DB	295	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	236
OY	9561	ATTTAAATATATCTGCGAAGATTGGAAAGGAAAGCTATATAAAAACGTCGGTAGATC	9620
DB	235	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	176
OY	9621	TAAATCAGTAGACTGTATACGATTAATATACATGGCATCTTCTGAAAAAGCATCAAA	9680
DB	175	AAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAAN	116
OY	9681	AAATATGACATATCAATCAAGAGAGAAAAAGGAAAAAA 9723	

Db	115	AAA	73
RESULT 9			
AC386893/c			
DEFINITION	AG386893	1519 bp	DNA
LOCUS	Mus musculus mojosinus DNA, clone:MSMg01-201E07.TJ, genomic survey		
ACCESSION	AG386893		
VERSION	AG386893.1	GI:47998098	
KEYWORDS	GSS.		
SOURCE	Mus musculus mojosinus (Japanese wild mouse)		
ORGANISM	Mus musculus mojosinus		
REFERENCE	Eulayotai, Metazoa; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 Aba,K., Noguchi,H., Tagawa,K., Yuzuriha,M., Toyoda,A., Kojima,T., Ezawa,K., Satou,N., Hattori,M., Sakaki,Y., Moriwaki,K. and Shiroishi,T.		
TITLE	Contribution of Asian mouse subspecies Mus musculus mojosinus to genomic constitution of strain C57BL/6J, as defined by BAC-end		
JOURNAL	Genome Res. 14 (12), 2439-2447 (2004)		
PUBMED	15574823		
REFERENCE	2 (bases 1 to 1519)		
AUTHORS	Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.		
TITLE	Direct Submission		
COMMENT	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Shinto-chou,Tsukumi-Ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:hattori@gs.c.riken.jp, URL:http://hnp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)		
	Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@crc.riken.jp).		
	Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan		
	phone: 81-298-35-9189, fax: 81-298-36-9199		
	e-mail: abe@crc.riken.jp		
	PRIMERS		
	Sequencing : TJ		
	LIBRARY		
	Vector : PBACe3.6		
	R.Site 1 : EcoRI		
	R.Site 2 : EcoRI.		
FEATURES	Location/Qualifiers		
source	1..1519		
	/organism="Mus musculus mojosinus"		
	/mol_type="genomic DNA"		
	/sub_species="mojosinus"		
	/db_xref="taxon:57486"		
	/clone="MSMg01-201E07.TJ"		
	/sex="male"		
	/issue_type="mixture of kidney and spleen"		
	/clone_lib="MSMg01 Mouse Male BAC Library"		
ORIGIN			
Query Match	1.0%; Score 96.8; DB 10; Length 1519;		
Best Local Similarity	38.8%; Pred. No. 1.4e-08;		
Matches 500; Conservative	0; Mismatches 789; Indels 0; Gaps 0;		
0y	8435	AAATATGATTACACAAAARATATTCAGAGCAATGCTTGCTTAATGATA	8494
Db	1512	ANNAN	1453
0y	8495	CAAGAATGAGCAATATAGAACTTTAGCGGTAGCGGTGTCATGCAAGAACAA	8554
Db	1452	NAAA	1393
0y	8555	ACAAAGATTAGAGATCAACAGTTAACTTCTACACTGTAAATGAGAGAAAGTAT	8614

```

Db      1392 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1333
Qy      8615 CTCACCTTGTCGCAAAAGCTTGGCTAAATAATGAAATTTATGAAATGTAAGAGACTG 8674
Db      1332 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1273
Qy      8675 GAGAGCCCTTAAGTCGAGCGGAAACAGACCGCTTGAAATTAATACAAAGACTACTACAG 8734
Db      1272 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1213
Qy      8735 GAGCATTGGTTGACGAAATTTGGGAAATTTGAGATTAATTAAGAAAGATTGCAAGATA 8794
Db      1212 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1153
Qy      8795 ATACGATTGTAAGATCAACGAGACGAAACCAAGAGAGTCTGTGCGAAAGAAATGTA 8854
Db      1152 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1093
Qy      8855 TTTCTGTGAATAATACATTTTCAGGGGAAACAAATCATCTGAAGATAAGCCAGAA 8914
Db      1092 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1033
Qy      8915 TTGTTGAAACCGGAAGTGTAAATGTGATGCTTTGATGATCACTTGATGTGATCTACAG 8974
Db      1032 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 973
Qy      8975 GAAAAAGTGCGCTATGCTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAAT 9034
Db      972 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 913
Qy      9035 AGAATAATGTAGAGCCAAATTCGAGACATGCTATTGTAGAACTGTGAAACCAAG 9094
Db      912 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 853
Qy      9095 AATATCAAGATTTCAGAGCAAAAGTAAATTTCTGGAAAGAGACCGCTCCAGCTG 9154
Db      852 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 793
Qy      9155 CAGCTGCAATTCGATGTACATTTCCATGAGATGATTAATAATTTGGCAAGC 9214
Db      792 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 733
Qy      9215 AGTATGATCTTCTCAATTAATTAACCAAAATTCAGAAATTAATTAATTTAGCATCA 9274
Db      732 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 673
Qy      9275 GTAGTGAATCGAATGTGATGTTTCATGGGGTGGCTGAGCAGAGGTGCGAGCCAAAG 9334
Db      672 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 613
Qy      9335 CGACAGTTAGTGAAGATCAATAATTAATGAATCTAATATGTTGATTTAGCAGAAAA 9394
Db      612 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 553
Qy      9395 TTAATAACAGAGGAAACATCAATGTATATGCCGAGATGATTAATAATTAATTAAGTA 9454
Db      552 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 493
Qy      9455 AGACAAATTTCTAAGGCTATTTGGCGATGCCAAAGTCACTGCTGCTGGCACTG 9514
Db      492 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 433
Qy      9515 CCCTATTGTAATAATTAAGTAATTTTAATGATCGAGAAATTTAAATAATTAATTC 9574
Db      432 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 373
Qy      9575 TGCAAGATTGAGAGGAAAGCTAATTAATAAAAGTCGATGATCTAATCAAGTAGACT 9634
Db      372 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 313
Qy      9635 GGTATACGATTAATTAATACATGCGATTCTTCTGAAAAAGCATCAAAAAATTGACATTC 9694
Db      312 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 253

```

```

Qy      9695 AATCAAGAGAGAGAAAAAGGAAAAAA 9723
Db      252 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 224

RESULT 10
DR124185
LOCUS
DEFINITION
pseudobscura cDNA clone M7 5', mRNA sequence.
ACCESSION
DR124185
VERSION
DR124185.1 GI:67840883
KEYWORDS
SOURCE
ORGANISM
Drosophila pseudoobscura
Drosophila pseudoobscura
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1691)
Richards,S., Liu,Y., Bertencourt,B.R., Hradecky,P., Letovsky,S.,
Nielsen,R., Thornton,K., Hubisz,M.J., Chen,R., Meisel,R.P.,
Coutome,O., Hua,S., Smith,M.A., Zhang,P., Liu,J., Bussemaker,H.J.,
van Batenburg,M.P., Howells,S.L., Scherer,S.E., Soderstrom,E.,
Matthews,B.B., Crosby,M.A., Schroeder,A.J., Ortiz-Barrileto,D.,
Rives,C.M., Metzker,M.L., Muzny,D.M., Scott,G., Steffen,D.,
Wheeler,D.A., Worley,K.C., Havlak,P., Durbin,K.J., Egan,A.,
Gill,R., Hume,J., Morgan,M.B., Miner,G., Hamilton,C., Huang,Y.,
Walton,L., Verdusco,D., Clerc-Blanchard,K.P., Dubchak,I.,
Noor,M.A., Anderson,M., White,K.P., Clark,A.G., Schaeffer,S.W.,
Gelbart,W., Weinstein,G.M. and Gibbs,R.A.
Comparative genome sequencing of Drosophila pseudoobscura:
chromosomal, gene, and cis-element evolution
Genome Res. 15 (1), 1-18 (2005)
15632085

JOURNAL
PUBMED
COMMENT
Contact: Stephen Richards
Human Genome Sequencing Center
Baylor College of Medicine
One Baylor Plaza, Houston, TX 77030, USA
Tel: 713-798-6667
Email: stephenr@bcm.tmc.edu
NCBI Trace Archive: 226706818
Insert length: 1750 Std Error: 0.25.
Location/Qualifiers
1..1691
/organism="Drosophila pseudoobscura"
/mol_type="mRNA"
/db_xref="taxon:7237"
/clone="M7"
/dev_stage="0-18h embryo"
/clone_lib="Drosophila pseudoobscura embryonic cDNA
library"
/note="Vector: pOTB7, Site_1: EcoRI; Site_2: XhoI, oligo
dt priming from poly A+ RNA, directionally cloned"

ORIGIN
Query Match 1.0%; Score 96.6; DB 8; Length 1691;
Best Local Similarity 42.1%; Pred. No. 1.6e-08;
Matches 549; Conservative 0; Mismatches 754; Indels 0; Gaps 0;

Qy      8421 AAGCAGATTATGAATTAATGATTACACAAAAAATTAATTTGAGAGTCAATGCTCT 8480
Db      364 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 423
Qy      8481 TGCTTTAAATGATACAAAGATGAAGCGAATTTAGATTTTAAAGCGTACCGGTGTGCA 8540
Db      424 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 483
Qy      8541 TGCAAGAGAAACAACAAGCATTTTACGATCAACAAGATTATCTTACAACTGTAAA 8600
Db      484 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 543
Qy      8601 TGAGAGAAACGTATCTCAACTTCGTGCAAAAGCTTTGGCTAAATAATTAAGAAA 8660

```

[illegible]

DB	1624	AAAAAAAAAGAAAGAAAAAGAAAAAGAAAAAGAAAGAAAAA	1666
RESULT 11			
LOCUS	CL118787		
DEFINITION	ISB1-72M23_T7.2 ISB1 Xenopus tropicalis genomic clone ISB1-72M23,		
ACCESSION	CL118787		
VERSION	CL118787.1		
KEYWORDS	GSF.		
SOURCE	Xenopus tropicalis (western clawed frog)		
ORGANISM	Xenopus tropicalis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana.		
AUTHORS	1 (bases 1 to 2001) Kramitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,		
TITLE	Mardi, E. and Wilson, R.		
JOURNAL	A physical map of the xenopus tropicalis genome		
COMMENT	Unpublished (2003) Contact: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@wustl.edu Insert Length: 75000 Std Error: 0.00 Seq primer: T7 TAATACGACTCATCTATATGGG Class: BAC ends High quality sequence start: 384 High quality sequence stop: 783. Location/Qualifiers 1..2001 /organism="Xenopus tropicalis" /mol_type="genomic DNA" /db_xref="taxon:8364" /clone="ISB1-72M23" /clone_id="ISB1" /note="Vector: pBelBAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"		
FEATURES			
SOURCE			
ORIGIN			
Query Match	1.0%; Score 96.2; DB 10; Length 2001;		
Best Local Similarity	42.8%; Pred. No. 2e-08;		
Matches 558; Conservative	0; Mismatches 736; Indels 11; Gaps 2;		
OY	8422 AAGCAAGTTATGAAATTAATGATTACACAAAAAATATATTTCAGAGTCATGCTCTT	8481	
DB	181 AA	240	
OY	8482 GCTTTAATGATACAAAGATGAAGCAATATAGATCTTTAGCGGTAGCCGTGTGAT	8541	
DB	241 NNAGGAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	300	
OY	8542 GCACAGAACAAACAAAGCATTTACAGATCAACAAAGTTAACTTCTACACTGTAAAT	8601	
DB	301 AAAAAAAAAAAAAAAAAAGAAATTAATATATTAATAAAGAAAAAAGAAAAAAGAA	360	
OY	8602 GGAGAAACGATCTCAACTTCGTCGCAAAAGCTTTGCTAAATAATGAATAATTGAAAT	8661	
DB	361 AAAAAAAAAAAAAAAAAATTAATAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	420	
OY	8662 GTAAAGAACTGAGAGAGCTTATGTCGAGCGGAAACAGCAGCCGTTGAATAATTATACA	8721	
DB	421 AAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA	480	
OY	8722 AAGGTACTACAGAGCATTTGGTTGCGAGAAATTTGGAAATTGGAGTTAAATTGAAACG	8781	
DB	481 AAGAAAAAAG	540	
OY	8782 ATTCAAGAGATTAATGAGATTGTAAGAGTCAACGAGACGGAACCAAGAGAGTCTTGTC	8841	
DB	541 AATAA	600	

Oy		8842	GGAAGAATGTAATTCTGTGAAAAAATCAATTTTCAGGGGAAACAAAATCATCATTTGAA	8901
Db		601	AA	660
Oy		8902	GATAAAGCCAGAAATTGTTGAAACCAGAAAGTGAATATGATGTGCTTGTAATGAACCTTGAT	8961
Db		661	AAGAAAGAAAGAAAAAAAAAAAA	720
Oy		8962	GTAGATCTACAGAAAGAGTGTGGCTATGTGTGAATTGGTATTGGAAATGTTGATGTA	9021
Db		721	AAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAATTAIAAAAAAAAAAAAAA	780
Oy		9022	AATATATGATTAAGAAAAATGTGAAAGCCAAATCCGAAACATCGTATTGTAGAAACT	9081
Db		781	AATTAATTAATTAAGAAAGAAAAA	840
Oy		9082	ACTGAAAGAACAAATATCAAGCATTTTCAAGAGCAAAGTAAATATCTTGAAAAAGGA	9141
Db		841	AAATTAATAAATTAABACAAAAAATTAABAAAAAAGAAAAAAGAAAAAAGAAAAA	900
Oy		9142	GACGCTGACGCTGCAGTGCATATCGAATGTACATTTCCAATGAGATGATTTAA	9201
Db		901	AAAAAGATTAATTAATTAIAAAAAAAAAAAAAAAAAAGAAAAAAGAAAAA	960
Oy		9202	AATTGGAAAGCAGATATGCACTTCTCAATTAATTAACCAAAATTCGAAAAATTAATTT	9261
Db		961	AAAAAAAAAATTAAGAAAAAACAATTAATTAIAAAAAAAAAAGAAAAATTAIAAAAA	1020
Oy		9262	ACTTAGCATCAAGTAGTGAATCGAATGTGTAATGTTTCATGGGGTGTGTAAGCAAGAGT	9321
Db		1021	AAAAAGAAATTAABAAAAAAGAAATTAAGAAATTAABAAAAAATTAABAAAAATGAAAAA	1077
Oy		9322	GCAGAGCCAAAGCGACGTTAGTGAAGAAATCAATTAATTAAGAACTAATTAATGTTGAT	9381
Db		1078	--AACAAAGAAAAAATAATTAIAAAAAAAAAAGAAAAAAGAAAGATTAIAAAAAATA	1135
Oy		9382	TTAGCAGAAAAAATTAABACAGAGGAAACATCAATGATTAATGCCGATATGATTAATAT	9441
Db		1136	ATGAGAGAAAGAAAAAAGAAAAAATAATTAAGTGAAGAAAAATTAATTAATGAAAAA	1195
Oy		9442	TATATATTAATTAAGTAACAATTTCTAAGCTATTGGCGATGGCAAAAGTCATGTGCAGCT	9501
Db		1196	AAAAAAAAAAGTATATCAAAAAAAAAAGAAAAAGAAAAAGAAATTAAGAAAAAGAAAA	1255
Oy		9502	GCTTGGCACTGCCATATTGAAAAAATGAAGTAAATTTAATTAATGCGATCGAGAA	9561
Db		1256	GAAAAAATAATTAIAAAAAAAAAAGAAAAAATAAAAAAAAAAAGAAATTAAGAAAGAAAAA	1315
Oy		9562	TTTTAAAAATATCTGCGAAGATTTGAGAGGAAAGCTATTAIAAAAAAGTGGTAGATCT	9621
Db		1316	AAAAAAAAATTA-----AAAAAAGAAAAAGAAATTAIAAAAAAAAAAAGAAATA	1369
Oy		9622	AATCAGTAGACTGTATATCGATTAATATATACATGCAATTTCTTGAAAAAAGCATACAAA	9681
Db		1370	AAAAATGAAAAAAAAAAAAAAAAAATTAAGAAAAACAGAGAGAAAGAAAAAAGAGGA	1429
Oy		9682	AAATTGACATATCAATCAAAGAGAGAAAAAGGAAAAAATGA	9726
Db		1430	AAAAAGATTAIAAAAAAAAAAAGAAAAAATATATGAIAAAAAAAAAAGAA	1474
RESULT 12				
CL066420				
LOCUS		1547 bp	DNA	linear
DEFINITION		CH216-108P4 Sp6.1 CH216 Xenopus tropicalis genomic clone		GSS 31-DEC-2003
VERSION		CH216-108P4_genomic survey sequence.		
KEYWORDS		CL066420.1 GI:40522333		
SOURCE		GSS.		
ORGANISM		Xenopus tropicalis (western clawed frog)		
		Bukerjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoae; Pipidae;		

```

QY 9021 AATATAATGATTAGAAAAATGTAGAGCCAAATCGAAGACATGCTATTGTAGAAC 9080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 771 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 830
QY 9081 TACTGAGAAACAAATATCAGCATTTACAGAGCAAAAGTAATATTTCTGAAAAGG 9140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 831 AAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAG 890
QY 9141 AGACGTGCGACGTGCGCTGCATATCGAATGTACACATTTCCAAATGAGATGTATTTA 9200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 891 AAAAAAAAAAGAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 950
QY 9201 AATTTGGCAAGACAGTATGCTCTTCTCAATTAATTAACCAAAATTCACAAAATATAT 9260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 951 AGAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1010
QY 9261 TACTTTAGCATCAAGTAGAATCGAATGTGAATGTTTCATGGGGTGGTGAAGCAAGG 9320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1011 AAAAAAAAAAGAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1070
QY 9321 TCGAGAGCCAAAGCCAGCACTTGTGTAAAGAAATCAATTAATTAAGAACTAATATGTTGA 9380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1071 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1130
QY 9381 TTATGAGAGAAAAATTAATAACAGAGGAAACATCAATGTATATGCGGATATGATATAA 9440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1131 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1190
QY 9441 TTATTAATTAAGTAGAACAATTTCTAAGGCTATTGGGATGCCAAAAGTCATGCTGACG 9500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1191 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1250
QY 9501 TCGTTCGCAACTGCCCATTTGAAATTAATGAATTAATTAATTAATTAATTAATTAATTA 9560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1251 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1310
QY 9561 ATTTAAAAATTAATCTGCAAGATTGGAAGGAAAGCTAATTAATAAAAGCTCGTAGATC 9620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1311 AGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1370
QY 9621 TAATCAGTAGACTGTATACGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1371 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1430
QY 9681 AAAATTTGACATTCATCAAGAGAGAGAAAAAGGAAAAATGA 9726
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1431 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 1476

```

```

RESULT 13
LOCUS CL081992 1784 bp DNA linear GSS 31-DBC-2003
DEFINITION CH216-165M9_RM4.1 CH216 Xenopus tropicalis genomic clone
ACCESSION CL081992
VERSION CL081992.1 GI:40537905
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
            Xenopodinae; Xenopus; Silurana.
            1 (bases 1 to 1784)
            Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submls@se@wustl.edu
            Insert Length: 175000 Std Error: 0.00
            Seq primer: RM4 ctcaaggcgatcgctgcgagc

```

Class: BAC ends  
High quality sequence start: 409  
High quality sequence stop: 503.  
Location/Qualifiers

FEATURES  
source  
1. 1784  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/strain="Nigerian frog"  
/db\_xref="taxon:8364"  
/clone="CH216-165M9"  
/sex="male"  
/cell\_line="Stock 248 F7A2, Inbred N7"  
/clone\_1ib="CH216"  
/note="Vector: pTARBAC2.1; CHOR1-216 Xenopus tropicalis  
BAC library"

## ORIGIN

Query Match 1.0%; Score 95.2; DB 10; Length 1784;  
Best Local Similarity 43.6%; Pred. No. 3.1e-08;  
Matches 515; Conservative 0; Mismatches 664; Indels 2; Gaps 2;

```

QY 8544 ACAGAGACAAACAAAGCATTTACGAGATCAACCAAGTTACTTACACCTGTAAATG 8603
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 AAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 361
QY 8604 AGCAAAACGATCTCACTTCGTCGCAAAAGCTTTGCTAAATAATGAATAATGAATAT 8663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 AGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 421
QY 8664 AAAAGGAATGAGAGAGCCCTTAGTGGAGCGGAAACAGACGCCGTTGAATAATTAACA 8723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 481
QY 8724 GAGTACTACAGAGCATTTGCTGACGAAATTTGGAAATTTGAGATTAATTAAGAAC 8783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 541
QY 8784 TCGAAGAGTAATTAATCAATTTGTAAGATCAACGAGACGAGAACCAAGAGAGCTTGT 8843
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 601
QY 8844 AAAAGAAATGATTTCTGTAAAAATTCATATTTACAGGGAAC-ACAAATCATCCATTGA 8902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 661
QY 8903 ATAAAGCCAGAAATTTGTTGAACCGAAGTGAATGTAGATGCTTTGAATGAACCTG 8962
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 721
QY 8963 TAGATCTACAGAAAAAAGTGTGCTATGTCGAAATTTGGTATTTGAATTTGATGTA 9022
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 781
QY 9023 ATATATGATTTAAGAAAAATGTGAAGCCAAATCGAAGACATCTAATTTGTGAACCT 9082
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 782 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 841
QY 9083 CTGAGAAAAACAAATTAATCAAGCATTTTACAGAGCAAAAGTAATATTTCTTGAAGAG 9142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 901
QY 9143 AGCTGACGCTGACGCTGCAATATGTAATGTACACTTTTCAATGAGATGATTTAAA 9202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 961
QY 9203 ATTTGGCAAGACAGTATGCACTTCTCAATTAATTAACAAAAATTCAAAATAATTAATTA 9262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 962 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1021
QY 9263 CTTTAGCATCAAGTAGAATGAATGTGAATGTTCAATGGGCTGGAAGCAAGAGCTG 9322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1022 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1081

```



QY	9333	AGCCGACATTGTGTTAAAGAAATTCAAATTAATGAAGCTAATATGTTGATTTTACAGAGAA	9332
Db	966	AA	10255
QY	9333	AATTAAAACAGAGGAGAAATCAATATGTTATATGCCCGATATGATATAAATTATATATAG	94522
Db	1026	AA	10685
QY	9433	TAAAGCAATTTCTAAGGCTATTGGCGATGCCAAAATCATCTGTCAGCTGCTTCGGCAAC	9512
Db	1086	AA	1145
QY	9513	TGCCACTATTGAAAAAAAAATGAAGTAAATTTTAATATCGATCCGAGATTTTAAATATA	95722
Db	1146	AA	12050
QY	9573	TCTGGCAAGATTGAAAGGAGAAAGCTAATATAAAAAACGTCCGTAGATCTAATCAGTGA	96322
Db	1206	AA	12655
QY	9633	CTGGTATTCGGATTAATTTACATGCGATTCTTCGAAAAAGCATACAAAAATTGACATA	96922
Db	1266	AA	13255
QY	9693	TCAATCAAGAGAGAGAGAAAAAGGAAAAATGA	9726
Db	1326	AA	1359
RESULT 15			
AG332292/c			
LOCUS	AG332292	1482 bp	DNA
DEFINITION	Mus musculus molossinus DNA, clone:MSWg01-122G19.TU, genomic survey	linear	GSS 18-DEC-2004
ACCESSION	AG332292		
VERSION	AG332292.1	GI:47905602	
KEYWORDS	GSS.		
SOURCE	Mus musculus molossinus (Japanese wild mouse)		
ORGANISM	Mus musculus molossinus		
REFERENCE	1 Abe, K., Noguchi, H., Tagawa, K., Yuzuriha, M., Toyoda, A., Kojima, T., Ezawa, K., Saitou, N., Hattori, M., Sakaki, Y., Moriaki, K. and Shiotohshi, T.		
AUTHORS	Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis		
TITLE	Genome Res. 14 (12), 2439-2447 (2004)		
JOURNAL	15574823		
PUBMED	2 (baae 1 to 1482)		
REFERENCE	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.		
AUTHORS	Direct Submission		
TITLE	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan		
JOURNAL	(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)		
COMMENT	Clones are derived from the mouse BAC library MSWg01. For BAC library availability, please contact Kuniya Abe (abe@rc.riken.jp). The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyada, Tsukuba, 305-0074 Japan		
	phone: 81-298-36-9189, fax: 81-298-36-9199		
	e-mail: abe@rc.riken.jp		
	PRIMERS		
	Sequencing : TJ		
	Library		
	Vector : pBac3.6		
	R.site 1 : EcoRI		
	R.site 2 : EcoRI		

FEATURES		Location/Qualifiers	
SOURCE		1..1482	
		/organism="Mus musculus moJossinus"	
		/mol_type="genomic DNA"	
		/sub_species="moJossinus"	
		/db_xref="taxon:57486"	
		/clone="MSMG01-122G19.1T"	
		/sex="male"	
		/issue_type="mixture of kidney and spleen"	
		/clone_lib="MSMG01 Mouse Male BAC Library"	
Query Match	1.0%; Score 93.9; DB 10; Length 1482;		
Base Local Similarity	40.1%; Pred. Mis.5.8e-08;		
Matches 434; Conservative	0; Mismatch:649; Indels	0; Gaps	0;
8641	AAAAATGAAATTTGTGGAATGTAAAGAACTGAGAGCCTTATGTCGAGCGAAACA	8700	
Db	1387 AA	13288	
Qy	8701 GCACCCGCTGGAATTTATACAAAGAGTACTACAGAGCAATTGTTGCAGAAATTGGAA	8760	
Db	1327 AA	12668	
Qy	8761 ATTGGAATTAATTAAGAAACGATTCGACAGATTAATCGATTGTAAAGTCAACGAGAC	88220	
Db	1267 AA	1208	
Qy	8821 GGAACCAAGAGGCTCTTGCGAAAGATGCTATTTCTGGAATAATACATTTCCAGG	8880	
Db	1207 AA	1148	
Qy	8881 GAAACAAATCATCATTTGAAGATTAAGCCAGAAATGTTGCAACCGAAGTGAATGTA	8940	
Db	1147 NNAANAAA	1088	
Qy	8941 GATGCTTTGAATGAAGCTTATGATGATCTACAGAGAAAAGTGTGCTATGTTGAATT	9000	
Db	1087 AA	1028	
Qy	9001 GGTATTGGAATGTTGATGTAAATATGATGTTAAGAAAATGTAAAGCCAAATCGCA	906	
Db	1027 AA	968	
Qy	9061 AGACATGCTATTGTAGAAACTACGAGAAAACAAGATATCAAGCATTTACAGAGCAAA	9120	
Db	967 AA	908	
Qy	9121 GTAAATATTCTTGGAAGAGAGAGCGCTGACGCTGACGCAATATGCAATGTACATT	9180	
Db	907 AA	848	
Qy	9181 TCCATGATGATGATTTAAATTTGGCAAGCAATGATCTTCTCAATTATTAACC	9240	
Db	847 AA	788	
Qy	9241 AAAAATTCAAAATAATTAATTAATCTTGACATCAAGTGTGAATGCAATGTCAT	9300	
Db	787 AA	728	
Qy	9301 GCGGTGCTGACAGCAAGGTGACAGAGCCAAAGCGACAGTTAGTGTAAAGATCAATA	9360	
Db	727 AA	668	
Qy	9361 AATGAACTAATTAATGTTGATTAGCAGAAATAATTAACAGAGGAAACATCAATGTA	9420	
Db	667 AA	608	
Qy	9421 TAGCGCGATATGATAAATAATTAATTAATAGTAAACAATCTAAGCGATTGCGGAT	9480	
Db	607 AA	548	
Qy	9481 GCCAAAAGTACGTGACGCTGCTTCGGCAACTGCCATTTGAAAAAAAAATGAATAAA	9540	

[illegible][illegible]

Db 1193 AATANNAAAAAAAAANNTTAAAAAAGANNAAAAAAAAAAAAAAAAAAAA 1252  
Qy 9603 AAAACGTCGTGATCTAATCAGGTAGACTGTATACGATTAATATCATGGATTC 9662  
Db 1253 NAAAAAANANANANATTAATAAATAAATAAATAAATAAATAAATAAATAA 1312  
Qy 9663 TTCTGAAAAAGCATACAAAAATTGACATATCAATCAAGAGAGAAAA 9714  
Db 1313 NTANAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1364

RESULT 17  
CG748014 1531 bp DNA linear GSS 24-OCT-2003  
LOCUS P041-4-B02.za Ppa EcorI BAC Library Pristionchus pacificus genomic,  
DEFINITION genomic survey sequence.  
ACCESSION CG748014  
VERSION CG748014.1 GI:37968940  
KEYWORDS GSS.  
SOURCE Pristionchus pacificus  
ORGANISM Pristionchus pacificus  
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;  
Neodiplogasteridae; Pristionchus.  
1 (bases 1 to 1531)  
Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,  
Bunfer,J., van der Meulen,M. and Sommer,R.J.  
An integrated physical and genetic map of the nematode Pristionchus  
pacificus  
Mol. Genet. Genomics 269 (5), 715-722 (2003)  
JOURNAL 12884007  
PUBMED Contact: Sommer RJ  
COMMENT Evolutionary Biology  
Max-Planck-Institute for Developmental Biology  
Spemannstr. 37-39, Tuebingen D-72076, Germany  
Tel.: 00497071601371  
Fax: 00497071601498  
Email: ralf.sommer@tuebingen.mpg.de  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..1531  
/organism="Pristionchus pacificus"  
/mol\_type="genomic DNA"  
/strain="California"  
/db\_xref="taxon:54126"  
/clone\_lib="Ppa EcorI BAC Library"  
/note="The library was generated by a partial digest of  
the genomic DNA with EcorI and cloning into the BAC  
vector."

ORIGIN  
Query Match 1.0%; Score 93.2; DB 10; Length 1531;  
Best Local Similarity 37.7%; Pred. No. 7.7e-08;  
Matches 491; Conservative 0; Mismatches 812; Indels 0; Gaps 0;  
Qy 8421 AAAAGCAAGTTATGAATTAATGATTAACAATAAATAATATTTCAGAACTCAATGCTCT 8480  
Db 180 AA 239  
Qy 8481 TGCTTAAATGATACAAAGATGAAGCATATAGATCTTTAGCGGTAGCCGTGTGCA 8540  
Db 240 AA 259  
Qy 8541 TGCAACAAGAACAAAGCATTTACGAGATCAAAAGTTAACTTTACAACTGTAAA 8600  
Db 300 NAAA 359  
Qy 8601 TGGAGGAACGTATCTCACTTCTGCGAAAGCTTTGGCTAAATAATGAATAATTATGAAA 8660  
Db 360 AA 419  
Qy 8661 TGTAAAGGAACGTGAGAGCCTTAGTCGAGCGGAACGACCGCTTGAAATTTATAC 8720  
Db 420 AA 479

Qy 8721 AAAAGTACTACAGAGCATTGGTTGCAGAAATTGGAAATTGGATTAATTGAAAC 8780  
Db 480 AA 539  
Qy 8781 GATTGCAAGATTAATACGATTGTGAAGTCAACGAGACGGAACCAAGAGAGCTTGT 8840  
Db 540 AA 599  
Qy 8841 CGAAAGAAATGGTATTTCTGTGAAAAATCAATTTACGGGAAACAAAATCATCTTGA 8900  
Db 600 AA 659  
Qy 8901 AGTAAAGCAGATTTGTTGGAACCGAGGTGTAATGTAGATGCTTGAATGACTTGA 8960  
Db 660 AA 719  
Qy 8961 TGTAGATCTACAGAAAGAGTGTGGCTATGATGGAATTGGTAAATGTTGATGT 9020  
Db 720 AA 779  
Qy 9021 AATTAATGTGATTAAGAAAAATGTAGAGCCAAATCGAAGACATGCTATTGTAGAAC 9080  
Db 780 AA 839  
Qy 9081 TACTGGAAGCAAGAAATATCAAGCATTTACAGAGCAAAAGTAAATATCTTGGAAAG 9140  
Db 840 AA 899  
Qy 9141 AGACGCTGACGTGACGTGCAATATCGAATGTACATTTCCAAATGAGATGATTTAA 9200  
Db 900 AA 959  
Qy 9201 AATTTGGCAAGCATGATGATCTTCTCAATTAATACCAAAATTCAAAAATTAAT 9260  
Db 960 AA 1019  
Qy 9261 TACTTAGCATCAAGTATGTAATCGAATGTGAATGTTCAATGGGTGCTGAAGCAGAG 9320  
Db 1020 AA 1079  
Qy 9321 TCAGAGGCAAGCAGCATGTTAGTGAAGATCAATTAATAGAACTAATATATGTTGA 9380  
Db 1080 AA 1139  
Qy 9381 TTTAGCAGAAAAATTAAAAACAGAGGAACATCAATGTATATGCCGATATGATAAAA 9440  
Db 1140 AA 1199  
Qy 9441 TTTAATTAATAGTAAGCAAAATTTAGGCTATTGGGATGCCAAAGTCATGCTGACG 9500  
Db 1200 AA 1259  
Qy 9501 TCGTTGGCACTGCCATATTGAATAAATAAGTAAATTTAATATGCGATCGCAG 9560  
Db 1260 AA 1319  
Qy 9561 ATTTAAATAATTAATCTGGCAAGATTGGAAGGAAGCTAAATAAAACGTCGTAGATC 9620  
Db 1320 AA 1379  
Qy 9621 TAATCAGGTAGCTGTTATACGATTAATATATATATATGCTTCTTGAAAAAGCAT 9680  
Db 1380 AA 1439  
Qy 9681 AAAATTAATGATCAATCAAGAGAGGAGAAAAAGGAAAAA 9723  
Db 1440 AA 1482

RESULT 18  
CL110653/c  
LOCUS CL110653 1594 bp DNA linear GSS 05-JAN-2004

DEFINITION ISB1-53P23\_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-53P23,  
genomic survey sequence.  
ACCESSION CL110653  
VERSION CL110653.1 GI:40604288  
KEYWORDS GSS.  
SOURCE Xenopus tropicalis (western clawed frog)  
ORGANISM Xenopus tropicalis  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
Xenopodinae; Xenopus; Silurana.  
1 (bases 1 to 1594)  
AUTHORS Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,  
Mardis,E. and Wilson,R.  
TITLE A physical map of the xenopus tropicalis genome  
JOURNAL Unpublished (2003)  
COMMENT Contact: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: subml@seion.wustl.edu  
Insert Length: 75000 Std Error: 0.00  
Seq primer: Sp6 ATTTAGTGACACTATAG  
Class: BAC ends  
High quality sequence start: 390  
High quality sequence stop: 470.  
Location/Qualifiers  
1..1594  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:8364"  
/clone="ISB1-53P23"  
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC  
Library Segment 1"

ORIGIN  
Query Match 1.0%; Score 93; DB 10; Length 1594;  
Best Local Similarity 43.1%; Pred. No. 8.6e-08;  
Matches 556; Conservative 0; Mismatches 725; Indels 8; Gaps 2;  
Library Segment 1"

8417 TACGAGAGAGATTGTAAGATAATGATTCACAAATAATATTGAGAAATGCAATG 8476  
1562 TAAAAAATAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATA 1503  
8477 CTCTTGCTTAAATGATACAAAGATGAGCGAATAGATCTTAGCGGTAGCCGTG 8536  
1502 AAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1443  
8537 TGCATGCACAGAGCAACAAAGCATTTACGAGATCAACAAAGTTACTTCTACAACTG 8596  
1442 TTAATTAACATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1383  
8597 TAAATGAGAGAAACGATCTCACTTCGTCAGAAAGCTTGGCTAATAATGAAATTATG 8656  
1382 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1330  
8657 GAAATGTAAAGAACTGAGAGAGCTTAGTCGAGACGGAACAGACGCGTTGAAATT 8716  
1329 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1270  
8717 ATACAAAGAGTACTACAGAGCATTTGGTCAGAAATTGGAAATTGAGATTAATTAG 8776  
1269 AATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1210  
8777 AAACGATTGCAAGAGATTAATGACATTGTGAAGCTCAACGAGACGAAAGAGGTC 8836  
1209 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1150  
8837 TTGTGCGAAGAGATGATTTCTGTGAATAATCAATTTCAAGCGGAAAGAAATCATCA 8896  
1149 ATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1090  
8897 TTGAAGATTAAGCCAGCAATTTGTTGGAACGGAAGTGAATGATGCTTTGAATGAAC 8956

Db 1089 AGAAAAAATCAAAAAAATTAATAATAATAATAATAATAATAATAATAATAATA 1030  
Qy 8957 TTGATGTAGATCTACAGAGAAAGTGGCTATGTGTGGAATTTGGAAATGTTG 9016  
Db 1029 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 970  
Qy 9017 ATGTAATTAATGTGATTAAGAAAAATGTGAAGCCAAATTCGAAAGACATGCTATTGTAG 9076  
Db 969 AATTAAAAAATCAAAAAAATTAATAATAATAATAATAATAATAATAATAATA 910  
Qy 9077 AAACATCTGGAAGAACAGAAATTCAGAGCTTTTCAAGAGCAAAAGTAATATTCTTGAA 9136  
Db 909 AAAAAAATTAACAA-CAAATTAATTAATAATAATAATAATAATAATAATAATA 851  
Qy 9137 AAGAGACGCTGACGCTGCAATTAATGATGATCACTTCCATGAGATGATGAT 9196  
Db 850 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 791  
Qy 9197 TTAATAATTTGCAAGAGATGATGATCTTCAATTAATAACAAATTCAAATAATA 9256  
Db 790 ATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 731  
Qy 9257 ATATTATCTTACATCAAGTATGATGATGATGATGATGATGATGATGATGAT 9316  
Db 730 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 671  
Qy 9317 GAGTGCAAGAGCCAAAGCAGTATGATGATGATGATGATGATGATGATGATG 9376  
Db 670 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 611  
Qy 9377 TTGATTTGACAGAAATAATAATAAGAGGAAACATCAATGATATGCGGATATGATA 9436  
Db 610 ATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 551  
Qy 9437 AAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 9496  
Db 550 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 491  
Qy 9497 CAGCTGCTTCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9556  
Db 490 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 431  
Qy 9557 GAGATTTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 9616  
Db 430 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 371  
Qy 9617 GATCTAATCAGTACTGTGATGATGATGATGATGATGATGATGATGATGAT 9676  
Db 370 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 311  
Qy 9677 ACAAATAATGACATTCATCAATCAAGACA 9705  
Db 310 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 282

RESULT 19  
CL118709  
LOCUS CL118709 1654 bp DNA linear GSS 05-JAN-2004  
DEFINITION ISB1-72115.T7.2 ISB1 Xenopus tropicalis genomic clone ISB1-72115,  
genomic survey sequence.  
ACCESSION CL118709  
VERSION CL118709.1 GI:40612344  
KEYWORDS GSS.  
SOURCE Xenopus tropicalis (western clawed frog)  
ORGANISM Xenopus tropicalis  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
Xenopodinae; Xenopus; Silurana.  
1 (bases 1 to 1654)  
AUTHORS Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,  
Mardis,E. and Wilson,R.  
TITLE A physical map of the xenopus tropicalis genome  
JOURNAL Unpublished (2003)

**COMMENT**

Contact: Richard C Wilson  
Genome Sequencing Center  
Emory University School of Medicine  
Email: [submissions@watsn.wustl.edu](mailto:submissions@watsn.wustl.edu)  
Insert Length: 75000 Std Error: 0.00  
Seq primer: T7 TATATAGACTCTACTATAGGG  
Class: BAC ends  
High quality sequence start: 468  
High quality sequence stop: 812.

## FEATURES

```
1..1654
/organism="Xenopus tropicalis"
mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-72115"
/clone_1lb="ISB1"
/note="vector: peltobac1; ISB-1 Xenopus tropicalis BAC
library Segment 1"
```

**ORIGIN**

[illegible][illegible]

LOCUS	CL119201	RESULT 20
DEFINITION	CL119201 1256 bp DNA linear	GSS 05-JAN-2004
ACCESSION	ISB1-76J12_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-76J12,	
VERSION	CL119201	
KEYWORDS	genomic survey sequence.	
SOURCE	CL119201.1 GI:40612836	
ORGANISM	GSS.	
REFERENCE	Xenopus tropicalis (western clawed frog)	
AUTHORS	Xenopus tropicalis	
TITLE	Eumylaria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
JOURNAL	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;	
COMMENT	Xenopodinae; Xenopus; Silurana.	
	1 (bases 1 to 1256)	
	Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,	
	Mardis,E. and Wilson,R.	
	A physical map of the xenopus tropicalis genome	
	Unpublished (2003)	
	Contact: Richard K Wilson	
	Genome Sequencing Center	
	Washington University School of Medicine	
	Email: submis1ions@wats1on.wu1st1.edu	
	Insert length: 75000 Std Error: 0.00	
	Seq primer: Sp6 ATTAGTGACACTATAG	
	Class: BAC ends	
	High quality sequence start: 233	
	High quality sequence stop: 402.	

## FEATURES

```
source
1. .1256
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-76J12"
```

ORIGIN

/clone\_lib="ISB1"  
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC  
Library Segment 1"

Query Match 0.9%; Score 91.8; DB 10; Length 1256;  
Best Local Similarity 37.4%; Pred. No. 1.4e-07;  
Matches 441; Conservative 0; Mismatches 739; Indels 0; Gaps 0;

QY 8544 ACAAGACAAAGCAATTTACGATCAACCAAGTTAACTTACAACTGTAATGG 8603  
DB 2 AA 61  
QY 8604 AGGAAACGTATCTCACTTCGTGCAAAAGCTTGGCTAAATAATGAAATGTAATG 8663  
DB 62 AA 121  
QY 8664 AAAAGGACCTGAGAGCCTTACTGAGCGGAAACGACGCCGTTGAAATTAACAA 8723  
DB 122 AA 181  
QY 8724 GAGTACTACAGAGCATTTGGTTCAGAGAAATTGGAAATTGAGATTAATTGAACGAT 8783  
DB 182 AA 241  
QY 8784 TGCAGAGATTAATACGATTTGTAAAGTCAACGAGACGAAACCAAGAGCTCTTGC 8843  
DB 242 AA 301  
QY 8844 AAAAGATGTTATTTCTGTGAAAAATCAATTTAGGGGAAACAAATCATCATGAGA 8903  
DB 302 AA 361  
QY 8904 TAAAGCAGATTTGTGAACCGAGGTGTAATGTAGATGCTTGAATGAATGATGAT 8963  
DB 362 AA 421  
QY 8964 AGATCTACAGAAAGAAAGTGTGTGATGATGATGATGATGATGATGATGATGAT 9023  
DB 422 AA 481  
QY 9024 TAAATGATTAAGAAAAATGTAGAGCCAAATCGAGACATGCTATTGTAGAACTAC 9083  
DB 482 AA 541  
QY 9084 TGGAAACAGAAATATCAAGCATTTTCAAGAGCAAAAGTAATCTTGGAAAGAGA 9143  
DB 542 AA 601  
QY 9144 CGCTGACGCTGACGCTGATATCGAATGACATTTTCCATGAGATGATTAATAA 9203  
DB 602 AA 661  
QY 9204 TTTGGCAAGCAGTATGATCTTCTCAATTAATACCAAAATTCACAAAATTAATATAC 9263  
DB 662 AA 721  
QY 9264 TTTAGCATCAAGTATGAGTATGCAATGTTCAATGGGCTGCTGAAGCAGAGTGC 9323  
DB 722 AA 781  
QY 9324 AGGAGCAAGCAGCAGTATGATTAAGATCAATTAATTAATTAATGATGATTT 9383  
DB 782 AA 841  
QY 9384 AGCAGAAAAATTTAAACAGAGGAAACATCAATGTATATCGCGATATGATTAATAA 9443  
DB 842 AA 901  
QY 9444 TAAATATGTAAGCAAAATTTCTAAGCTATTGCGGATGCGAAAGTCATGCTGACATGC 9503  
DB 902 AA 961

QY 9504 TTGCGCACTGCGACTATTGAAAAAATGAACTAAATTTAATTAATGCGATCGAGATT 9563  
DB 962 AA 1021  
QY 9564 TAAATAATTAATCTGCGAAGTTGGAAGGAAAGCTAATTAATAAGCTGTGATCTTA 9623  
DB 1022 AA 1081  
QY 9624 TCAGGTAGCTGTGTACGATTAATTAATCATGAGCTTCTGCAAAAGCATACAAA 9683  
DB 1082 AA 1141  
QY 9684 ATTGACATATCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9723  
DB 1142 AA 1181

RESULT 21  
CL076558 1485 bp DNA linear GSS 31-DEC-2003  
LOCUS CH216-139010 Sp6.1 CH216 Xenopus tropicalis genomic clone  
DEFINITION CH216-139010, genomic survey sequence.  
ACCESSION CL076558  
VERSION CL076558.1 GI:40532471  
KEYWORDS GSS.  
SOURCE Xenopus tropicalis (western clawed frog)  
ORGANISM Xenopus tropicalis  
Eukaryote; Metazoa; Chordata; Vertebrata; Ruteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;  
Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 1485)  
Kremliutzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,  
Mardis, E. and Wilson, R.  
A physical map of the xenopus tropicalis genome  
Unpublished (2003)  
CONTACT: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@wustl.edu  
Insert Length: 175000 Std Error: 0.00  
Seq primer: Sp6 ATTTAGGTGACACTATAG  
Clase: BAC ends  
High quality sequence start: 572  
High quality sequence stop: 635.  
Location/Qualifiers

FEATURES  
source 1. 1485  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/strain="Nigerian frog"  
/db\_xref="taxon:8364"  
/clone="CH216-139010"  
/sex="male"  
/cell\_line="Stock 248 F7A2, inbred N7"  
/clone\_lib="CH216"  
/note="Vector: pTAKBAC2.1; CHORI-216 Xenopus tropicalis  
BAC library"

ORIGIN

Query Match 0.9%; Score 91.8; DB 10; Length 1485;  
Best Local Similarity 42.8%; Pred. No. 1.5e-07;  
Matches 561; Conservative 0; Mismatches 745; Indels 4; Gaps 2;

QY 8421 AAAGCAAGTTATGAAATTAATGATTAACAAAAATTTATTTGAGAGTCAATGCTCT 8480  
DB 111 AA 170  
QY 8481 TCGTTAAATGATCAAGAAATGAAGCAATATTAATCTTTAGCGGTAGCCGTGTGCA 8540  
DB 171 AA 230  
QY 8541 TGCACAGACAAACAAAGCATTTTACGAGATCAACAGTTTACTTCTACAACTGTAA 8600  
DB 231 GATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAGAGAGAGAGAGAGAGAGAGAGAG 290





```

Qy 9133 GGAAGAGAGCGTGCAGTGCATGCAATATCGAATGTACATTTCCATGAGATG 9192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1044 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 985
Qy 9193 GATATTAATTTGGCAAGCATGTGCTTCTCAATTAATTAACCAAAATTCAAA 9252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 984 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 925
Qy 9253 AATTAATTTACTTGTAGCATCAAGTAGTAATCGAATGTGAATGTTCAATGGGCTGAA 9312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 924 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 865
Qy 9313 GCAAGAGTGCAGAGCGCAAGCGACGTTAGTGAATGAATCAATTAATGACATTAAT 9372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 864 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 805
Qy 9373 AATGTTGATTTAGCAGAAAAATTTAAAAACAGAGAAACATCAATGTATGCGGATAT 9432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 804 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 745
Qy 9433 GATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 744 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 685
Qy 9493 GCTGCGAGCTGCTCGGCACTGCACTATGAAATGAATGAATTAATTAATTCG 9552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 684 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 625
Qy 9553 ATCCGAGATTTTAAAAATATCTGCAAGATGGAAGGAAAGCTAATTAATAAAGCTCG 9612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 624 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 565
Qy 9613 GTTGAATCTTAATCAAGTAGCTGTATACGATTAATTAATCATGCACTTCTTGA 9672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 564 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 505
Qy 9673 GCATCAAAAAATTTGACATATCAATCAAGAGGAAAGGAAAAA 9723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 504 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 454

```

RESULT 24  
AJ591978/c  
LOCUS  
DEFINITION  
Arabidopsis thaliana T-DNA flanking sequence, right border, clone 598D01, genomic survey sequence.  
ACCESSION  
AJ591978  
VERSION  
AJ591978.1 GI:37941602  
KEYWORDS  
GSS, right border; T-DNA flanking sequence.  
SOURCE  
Arabidopsis thaliana (thale cress)  
ORGANISM  
Arabidopsis thaliana  
REFERENCE  
1 Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Crnaud, C., Dehose, R., Pelletier, G., Lepoint, L., Caboche, M. and Lecharny, A.  
T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites  
EMBO Rep. 3 (12), 1152-1157 (2002)

JOURNAL  
PUBMED  
12446565  
2 (bases 1 to 1453)  
REFERENCE  
AUTHORS  
Balzergue, S.  
TITLE  
Direct Submision  
JOURNAL  
Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE  
COMMENT  
PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a

graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

FEATURES  
source  
1..1453  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"  
/clone="598D01"  
/clone\_1ib="Arabidopsis thaliana T-DNA insertion lines"  
/ecotype="Masliewskja"  
1..1453  
/note="T-DNA flanking sequence  
right border"

ORIGIN  
Query Match 0.9%; Score 91.4; DB 10; Length 1453;  
Best Local Similarity 37.9%; Pred. No. 1.8e-07;  
Matches 494; Conservative 0; Mismatches 809; Indels 0; Gaps 0;

```

Qy 8421 AAAGCAAGATTGAAATTAATGATTACAAAAAATATATTTCAGAGTCAATGCTCT 8480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1420 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1361
Qy 8481 TGCTTTAAATGATACAAAGATGACGAAATATGAATCTTTAGCGGTACCGGTGCA 8540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1360 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1301
Qy 8541 TGCACAAGAACAAACAAAGCATTTACGAGTCAACCAATTAATCTTACATCTGTA 8600
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1300 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1241
Qy 8601 TCGAGGAAACGATCTCAACTTCGTCGCAAAAGCTTGGCTTAAATGAATTTATGGA 8660
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1240 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1181
Qy 8661 TGTAAAGAACTGAGAGACCTTACTGAGCGGAAACGACCGCTTGAATTTATAC 8720
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1180 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1121
Qy 8721 AAAGACTACTACAGAGCATGTTGTCAGAGAAATTTGGCAATTTGATTAATTAAGAA 8780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1061
Qy 8781 GATTCAAGAGATTAATACGATTGTGAAGCTCAACGAGACGAAACCAAGAGATCTTGT 8840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1001
Qy 8841 CGGAAGAAATGATATTTCTGTGAAAAATCAATTTCAAGGGAACAAATCATCATTTGA 8900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 941
Qy 8901 AGATTAAGCCAGATTTGTTGGAACCGAGAGTGAATGTGAATCTTTGAATCTTGA 8960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 881
Qy 8961 TGTAGATCTTCAAGAGAAAGTGTGCTATGTGTGAATTTGTATTTGAATTTGATGT 9020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 880 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821
Qy 9021 AATTAATGTGATTAGAAAAATGTAGAGCCAAATCGAAGACATGTTATTTAGAAC 9080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
Qy 9081 TATGGAAGAAACAAGATTAATCAAGCATTTCAAGAGGAAAGTAAATTTCTTGAAGAG 9140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
Qy 9141 AGACGCTGAGCTGCAGCTGCATATCGAATGTACATTTCCATATGAGTGAATTTAA 9200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641

```

[illegible]

RESULT	25
LOCUS	CL109409
DEFINITION	CL109409 1829 bp DNA linear GSS 05-JAN-2004 ISB1-51H21_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-51H21, genomic survey sequence.
ACCESSION	CL109409
VERSION	CL109409.1 GI:4603044
KEYWORDS	GSS.
SOURCE	Xenopus tropicalis (western clawed frog)
ORGANISM	Xenopus tropicalis
REFERENCE	Bukacinska, Metazoa; Chordata; Craniata; Vertebrate; Buteleaeom; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae; Xenopodine; Xenopus; Silurane. 1 (bases 1 to 1829) Kremelicki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R. A physical map of the xenopus tropicalis genome Unpublished (2003) Contact: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: submissions@watson.wustl.edu Insert Length: 7500 Std Error: 0.00 Seq primer: Sp6 ATTAGGTGACATCATAG Class: BAC ends High quality sequence start: 856 High quality sequence stop: 936.
TITLE	
JOURNAL	
COMMENT	

```

FEATURES
    source
        location/Qualifiers
            1..1829
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-51H21"
                /clone_1jb="ISB1"
                /note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC"

```

ORIGIN	Library Segment 1"
Query Match	0.9%; Score 91.4; DB 10; Length 1929;
Best Local Similarity	42.4%; Pred. No. 1.9e-07;
Matches 568; Conservative	0; Mismatches 763; Indels 10; Gaps 1;
Qy	8383 ACAGCACAAGATATCTAGATCCGTAGTTCGTTTACGAAAGCAAGATTATGAAATATAT 8442
Db	80 AGAAGAAAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 139
Qy	8443 GATTACACAAAAATATATTTTCAGAGTCATGCTCTTGCTTTAATGATACAAAGAT 8502
Db	140 ATAAAAACAAAAAGAAAAAGAAAGAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 199
Qy	8503 GAAGCGAATATAGAACTTTAGCCGTAGCCGTGTGCATGCACAGAGACAAACAAAGCA 8562
Db	200 AAAAAAAGATAGAACCAACAAAAAGAAAAAGAAAAAGCAAGAAAAATATAAAAA 259
Qy	8563 TTTACGAGATCAACCAAGTTTACTTCACATGTAAATGAGAGAAACGATCTCACTT 8622
Db	260 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 319
Qy	8623 CGTCAAAAGCTTGCTTAAATATGAAATTTATGAAATGTAAAGAGAACTGAGAGACC 8682
Db	320 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 379
Qy	8683 TTACTCGAGCGGAAACAGACCCGTTGAAATTTATCAAGAGTACTACAGAGACTTG 8742
Db	380 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAATTAAGAGAA 439
Qy	8743 GTTCAGAGAAATTTGGGAAATTTGAGATTAATTTAGAAACGATTGCAGAGATTAATGATT 8802
Db	440 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 499
Qy	8803 GTAGAGTCACCGGAGCGAACCAAGAGAGTCTGTGGAAGAAATGGTATTTCTGTG 8862
Db	500 GAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 559
Qy	8863 AAAATATCAATTTTCAGGGGAAACAAATTCATCTTGAAGATTAAGCCAGAAATTTGTGA 8922
Db	560 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 619
Qy	8923 ACCGGAAGTGAATGTAGATCTTTGAATGAATCTGATGTAGATCTACAGAGAAAAAGT 8982
Db	620 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 679
Qy	8983 GGTGCTATGTGTGAATTTGTAATTTGAAATGTTGATGTAATTAATGTGATTAAGAAAAAT 9042
Db	680 AAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 739
Qy	9043 GTAGAACCAAAATCGAGAGCATGCTATTGTAGAACTACTGAAAAACAGATATCA 9102
Db	740 GAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 799
Qy	9103 GCATTTACAGAGCAAAAGTAATATTTCTTGAAAGAGACGCTGCAGCTGCAG 9162
Db	800 AACAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 849
Qy	9163 ATATCGAATGTACATTTCCATGGAATGTGATTTTAAATTTGGCAAGCGATATGCA 9222
Db	850 GAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 909
Qy	9223 TCTTCTCAATTAATTAACCAAAATTTCAAAATATATATTTACTTTAGCATCAAGTGTGA 9282
Db	910 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 969
Qy	9283 TCGAATGTGAATGTCTACGCGGTGCTGAAGCAAGAGTGCAGAGCCAAAGCAGACTT 9342
Db	970 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 1029
Qy	9343 AGTGTAAAGATCAATTAATAGAACTTAATTAATGTGATTTAGCAGAAAAATTAACA 9402

Db	1030	AAAAAAAAAAAAAAAAAAAAAAAAAGGAAAAAAGGAAAAAAAAAAAAAAAA
Qy	9403	GAGGAAACATCATGTATATGCCGATATGATAAATTATATATATAGTACAAAT
Db	1090	AAAAGAAAAAGAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGGA
Qy	9463	TCTAAGCGTATTCGGGATGCCAAAGTCATGCTGCCTCCCAACTCCTATT
Db	1150	AAAAAGAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAA
Qy	9523	GAUUAUUATGAATTAATTTATATATGATCCAGATTTTAAATATATCTGCCAGA
Db	1210	AAGAAAAAGAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAA
Qy	9583	TTGGAAGGGAACCTATATATAAAAAGCTGCGTAGATCTATACAGTAGACTGTATACG
Db	1270	AAGAAAAAGAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAA
Qy	9643	GATAATATATCATGCGCATTTCTTGAAAAAGATACAATAATTTGACATATCATCAAG
Db	1330	AAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAA
Qy	9703	AGAGGAGAAAAAGGAAAAA
Db	1390	AAAAAGAAAAAGAAAAAGAAAAA
RESULT 26		
CL082000/c		
LOCUS	CL082000	1981 bp DNA linear GSS 31-DEC-2003
DEFINITION	CH216-165P18_RM4.1 CH216 Xenopus tropicalis genomic clone	
VERSION	CH216-165P18, genomic survey sequence.	
KEYWORDS	CL082000	
SOURCE	CL082000.1 GI:40537913	
ORGANISM	Xenopus tropicalis (western clawed frog)	
REFERENCE	Xenopus tropicalis	
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana. Krenitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardle, B. and Wilson, R. A physical map of the xenopus tropicalis genome Unpublished (2003) Contract: Richard K Wilson Genome Sequencing Center Washington University School of Medicine Email: rdbm1elions@watson.wustl.edu Insert Length: 175000 Std Error: 0.00 Seq primer: RM4 ctcaaggcgcgctcgagc Class: BAC ends High quality sequence start: 265 High quality sequence stop: 497. Location/Qualifiers	
FEATURES	source	
ORIGIN		
Query Match	0.9%	Score 91.2; DB 10; Length 1981;
Best Local Similarity	42.4%	Pred. No. 2,1e-07;
Matches 626; Conservative	0; Mismatches 843; Indels	9; Gaps 2.;
8246 AAGGAAATTTGTCAGAACAAATGCATTGAATGCAATTTCTTAAAGTACGAGGTTTGATG		8305



REFERENCE	TITLE	JOURNAL	COMMENT
AG289820/C	AG289820	1401 bp DNA	linear GSS 18-DEC-2004
LOCUS	Mus musculus molossinus DNA, clone:MSMg01-064E24.TJ, genomic survey		
DEFINITION	sequence.		
ACCESSION	AG289820	GI:47862697	
VERSION	AG289820.1		
KEYWORDS	GSS.		
SOURCE	Mus musculus molossinus (Japanese wild mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 Abe, K., Noguchi, H., Tagawa, K., Yuzurika, M., Toyoda, A., Kojima, T., Ezawa, K., Saito, N., Hattori, M., Sakaki, Y., Moriaki, K. and Shirosaki, T.		
TITLE	Contribution of Asian mouse subspecies Mus musculus molossinus to genomic constitution of strain C57BL/6J, as defined by BAC-end sequence-SNP analysis		
JOURNAL	Genome Res. 14 (12), 2439-2447 (2004)		
PUBMED	15574823		
REFERENCE	2 (bases 1 to 1401)		
AUTHORS	Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suenho-cho, Tsukumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)		
COMMENT	Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan phone: 81-298-36-9189, fax: 81-298-36-9199 e-mail: abe@rtc.riken.jp		
PRIMERS			
LIBRARY	Sequencing : TJ		
VECTOR	: pBAC3.6		
R.Site 1	: EcoRI		
R.Site 2	: EcoRI.		
Location/Qualifiers	1..1401		
FEATURES	/organism="Mus musculus molossinus"		
source	/mol_type="genomic DNA"		
	/sub_species="molossinus"		

[illegible]

Db 495 NNNNNNNNTGGGGGGGGGAAAAAAAAAAAAAAAANNNNNNNNNANNNACMAAAAAAAAAAAAA 436  
QY 9601 AAAAAACGTCGTAGACTCTATCTAGTACTGCTATACGATTAATATATCATGCGAT 9660  
Db 435 AA 376  
QY 9661 TCTTCTGAAAAAGCATACAAAAAATTGACATATCAATCAAGAGAGAGAAAAAGGAAA 9720  
Db 375 AAA 316  
QY 9721 AA 9722  
Db 315 AA 314

RESULT 29  
AJ592026/c 1407 bp DNA 1linear GSS 15-JAN-2004  
LOCUS Arabidopsis thaliana T-DNA flanking sequence, right border, clone  
DEFINITION 600D05, genomic survey sequence.  
ACCESSION AJ592026  
VERSION AJ592026.1 GI:37941650  
KEYWORDS GSS; right border; T-DNA flanking sequence.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1  
Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,  
Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,  
Lepoint, L., Caboche, M. and Lecharny, A.  
T-DNA integration into the Arabidopsis genome depends on sequences  
of pre-insertion sites  
EMBO Rep. 3 (12), 1152-1157 (2002)  
12446565  
2 (bases 1 to 1407)

JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue  
Gaston Cremieux, 91057 Evry cedex, FRANCE  
Direct Submission  
PCR was performed on DNA from transformants of Arabidopsis thaliana  
plants from INRA (Versailles). The DNA fragment(s) resulting from  
the PCR were directly sequenced from the left or the right border  
to determine the genomic sequence flanking the insertion. T-DNA  
derived sequences were removed. Information to order the  
corresponding mutant line and a link to a database providing a  
graphical display of the insertion site are available at  
http://dbsgap.versailles.inra.fr/publiclines/. This sequence has  
been generated in the framework of the French plant genomics  
program 'Genoplante' (http://www.genoplante.com and  
http://genoplante-info.inbio.gen.fr).  
Location/Qualifiers  
1. 1407  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:3702"  
/clone="600D05"  
/clone\_1lb="Arabidopsis thaliana T-DNA insertion lines"  
/ecotype="Massiliae1a"  
1. 1407  
/note="T-DNA flanking sequence  
right border"

ORIGIN  
Query Match 0.94; Score 90.2; DB 10; Length 1407;  
Best Local Similarity 37.54; Pred. No. 3.1e-07;  
Matches 488; Conservative 0; Mismatches 815; Indels 0; Gaps 0;  
QY 8421 AAGAGAGATTAGAAATTAATGATTAACAAAAAATATATTTGGAAGTCAATGCTCT 8480  
Db 1330 AAA 1271

QY 8481 TCGTTAAATGATACAAAGAAATGAACGATATAGATCTTTAGCGGTAGCCGTGTGCA 8540  
Db 1270 AAA 1211  
QY 8541 TGCACAGAGACAAACAAAGCATTTACGAGATCAACAGTTTACTTCTACACTGTAAA 8600  
Db 1210 AAA 1151  
QY 8601 TCGAGAAACGTACTCAACTCTGTCGAAAGCTTTGGCTAAATAATGAAAATTATGAAA 8660  
Db 1150 AAA 1091  
QY 8661 TGTAAAGAACTGAGAGACCTTAGTCGAGCGGAAACAGCGCTTTGAAATTTATAC 8720  
Db 1090 AAA 1031  
QY 8721 AAGAGTACTACGAGACATTGTTGTCGAGAAATTGGGAAATTGGAGATTAATTGAAAC 8780  
Db 1030 AAA 971  
QY 8781 GATTGCAAGATTAATACGATTGTAAAGTCAACGAGACGGAACCAAGAGGTCTGT 8840  
Db 970 AAA 911  
QY 8841 CGAAAGAAATGTTATTTCTGTGAAAAATACATTTCAAGGGAACAAATCATCATTTGA 8900  
Db 910 NAAAN 851  
QY 8901 AGATAAGCCGAAATTGTTGAAACCGAAAGTAAATGTAGTCTTGAATGAACCTTGA 8960  
Db 850 AAN 791  
QY 8961 TGTAGATCTACAGAAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9020  
Db 790 AAN 731  
QY 9021 AATATATGATTAAGAAAAATGTAGAAAGCCAAATCGAGACATGCTATTGTGAAC 9080  
Db 730 AAN 671  
QY 9081 TACTGAAACCAAGATATCAAGCTTTTCAAGACCAAAAGTAAATTTCTTGGAAAAG 9140  
Db 670 AAN 611  
QY 9141 AGACGCTGAGCTGAGCTGCAATTCGATGTACATTTCCAAATGAGATGATATTAA 9200  
Db 610 AAN 551  
QY 9201 AATTTGCAAGCAGTATGCACTTCTCAATTAATACCAAAATTCAAAAATTAATAT 9260  
Db 550 AAN 491  
QY 9261 TACTTACGATTAAGTATGGAATCGAATGTGATGTTCTAGGGGTGCTGAAGCAAGAG 9320  
Db 490 AAN 431  
QY 9321 TGCAGAGCAAGGAGCAGTTAGTAAAGATCAATTAATTAATAGAACTTAATATGTTGA 9380  
Db 430 AAN 371  
QY 9381 TTTACAGAAAAATTAACCAAGGAAACATCAATGATATATGCGGATATGATAAAA 9440  
Db 370 AAN 311  
QY 9441 TTTATATTAAGTAAAGCAATTTCTAAGCTATTGCGATGCCAAAGTCATGCTGCAGC 9500  
Db 310 AAN 251  
QY 9501 TGCCTGCGCACTGCACTATTGAAGAAAAATGAATTAATTAATTAATGATCCGAGA 9560  
Db 250 AAN 191  
QY 9561 ATTAAATAATCTGCAAGATTGGAAGGAAAGCTAATAAATAAAGCTCGGTAGATC 9620





**TITLE**  
Gelbart, W., Weinstock, G.M. and Gibbs, R.A.  
Comparative genome sequencing of *Drosophila pseudoobscura*:  
Chromosomal, gene, and cis-element evolution  
**JOURNAL**  
Genome Res. 15 (1), 1-18 (2005)  
**PUBMED**  
15632085  
**COMMENT**  
Contact: Stephen Richards  
Human Genome Sequencing Center  
Baylor College of Medicine  
One Baylor Plaza, Houston, TX 77030, USA  
Tel: 713-798-6667  
Email: stephenr@bcm.tmc.edu  
NCBI Trace Archive: 226717045  
Insert Length: 1750 Std Error: 0.25.

**FEATURES**  
source  
1.1932  
location/Qualifiers  
/organism="Drosophila pseudoobscura"  
/mol\_type="mRNA"  
/db\_xref="taxon:7237"  
/clone="113"  
/dev\_stage="0-18h embryos"  
/clone\_lib="Drosophila pseudoobscura embryonic cDNA  
library"  
/note="Vector: pOT87; Site 1: EcoRI; site 2: XhoI; oligo  
dc priming from poly A+ RNA, directionally cloned"

**ORIGIN**

Query Match 0.9%; Score 89.2; DB 8; Length 1932;  
Best Local Similarity 41.4%; Pred. No. 5.3e-07;  
Matches 538; Conservative 0; Mismatches 763; Indels 0; Gaps 0;

8422 AAGCAAGATTGAAATATGATTCACAAAAAATATTTTCAGAGTCATGCTCT 8481  
264 AAAN 323  
8482 GCTTAAATGATCAAGATGAGCAATATGAACTTTACGGTAGCCGTGTCAT 8541  
324 AAAAAACAACAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 383  
8542 GCACAGAACAAACAGATTGACAGATCAACAGTTACTTACACTGTAAAT 8601  
384 AAAAAAGAGAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 443  
8602 GAGAGAAAGTATCTCACTTGTGCAAAAGCTTTGGCTTAAATGAAATTGAAAT 8661  
444 AAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 503  
8662 GTTAAAGAACTGAGAGCCTTAGTCGAGCGGAAACAGACCCGTTGAAATTATCA 8721  
504 AA 563  
8722 AAGAGTACTACAGACATTGTTGACAGAAATTGGAAATTGAGATTAATTGAAACG 8781  
564 GAAAAAAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 623  
8782 ATTGCAAGAGATATAGATTGATGAGTCAAGAGACGAGCAACAAAGAGCTTGTG 8841  
624 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 683  
8842 GGAAGAAATGTTCTGTGAAAAATACATTTTCAGGGGAAACAAATCATCTTGA 8901  
684 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 743  
8902 GATTAAGCCAGATTGTTGACCGGAGGTGAATGTAGATCTTTGAATGAATTGAT 8961  
744 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 803  
8962 GTAGATCTACAGAGAAAGTGTGCTATGTGGAATTGTGTAATTGTAATGTA 9021  
804 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 863  
9022 AATTAATGATTAAGAAAAATGTAGAACCAAAATCGAAGACATGCTATTGTGAAC 9081  
864 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 923

9082 ACTGGAACACAGAAATATCAAGATTATCAAGAGCAAAAGTAAATTTCTTGAAAAAGA 9141  
924 AAGACG 983  
9142 GACGCTGACAGCTGACAGCTCAATATCGAATGTACATTTCCAAATGAGATATATAA 9201  
984 AAGAGAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1043  
9202 AATTGGCAAGCAGTATGATCTTCTCAATTATTAACCAAAATTCAAAAATTAATT 9261  
1044 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1103  
9262 ACTTTAGCTACAGTATGTAATCGAATGTGAATGTTTCAGGGGTGCTAAGCAAGAGT 9321  
1104 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1163  
9322 GAGAGAGCCAAAGCGACAGTATGTTAAGAAATCAAAATTAATTAATGTTGAT 9381  
1164 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1223  
9382 TTAGCAGAAAAATTAACAGAGGAAACATCATGTATGTATGCCGATATGATAAAT 9441  
1224 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1283  
9442 TATTAATTAATGTAAGCAAAATTCATAGCTATTCGCAATCCAAAGTCATGAGCT 9501  
1284 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1343  
9502 GCTTGGCAACTGCCACTTATGAAAAAATGAAGTAAATTTATATATGATCCAGAA 9561  
1344 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1403  
9562 TTTAAATTAATCTGGCAAGATTGAGGAAAGCTAATTAATAAAGCTGCGTAGACT 9621  
1404 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1463  
9622 AATCAGTATGATGCTGATACGATTAATATACATGCTCTTCGAAAAAGCATCAAA 9681  
1464 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1523  
9682 AATTAATGATATCAATCAAGAGAGAGAAAAAGGAAAAA 9722  
1524 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAA 1564

**RESULT 33**  
CC187638 1434 bp DNA linear GSS 08-MAY-2003  
**LOCUS**  
CH261-98P15\_Sp6.1 CH261 Gallus gallus genomic clone CH261-98P15,  
genomic survey sequence.  
**DEFINITION**  
CC187638  
**ACCESSION**  
CC187638  
**VERSION**  
CC187638.1 GI:30431538  
**KEYWORDS**  
GSS.  
**SOURCE**  
Gallus gallus (chicken)  
**ORGANISM**  
Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 1434)  
Krentlitz, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,  
Warren, W., Graves, T., Mardis, E. and Wilson, R.  
Gallus gallus BAC End Reads  
Unpublished (2003)  
Contact: Richard K. Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submissions@atson.wustl.edu  
Insert Length: 18200 Std Error: 0.00  
Seq primer: Sp6 ATTAGTGACACTATAG  
Class: BAC ends  
High quality sequence start: 422  
High quality sequence stop: 487.

**TITLE**  
JOURNAL  
**COMMENT**

# FEATURES Source

Location/Qualifiers  
1. 1434  
/organism="Gallus gallus"  
/mol\_type="genomic DNA"  
/strain="Red Jungle Fowl"  
/db\_xref="taxon:9031"  
/clone="CH261-98P15"  
/sex="female"  
/cell\_line="UCD001, Indred 256"  
/clone\_1lb="CH261"  
/note="Vector: pTARBAC2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
CH261 Female Chicken library - for library and clone  
ordering information: http://www.chori.org/bacpac"

## ORIGIN

Query Match 0 %; Score 88.6; DB 9; Length 1434;  
Best Local Similarity 40.8%; Pred. No. 6.6e-07;  
Matches 447; Conservative 0; Mismatches 648; Indels 1; Gaps 1;

```

QY 8628 AAAAGCTTGGCTAAATGAAATATGAAAGAACTGAGAGCTTAGT 8687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 13 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8688 CCGAGCGAAGACGACGCCGTTGAAATATACAAAGTACTACAGAGCAATTGGTGC 8747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 73 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8748 AGGAATTTGGGAAATTTGAGATTAATTGAAACGATTGCAAGAGATTAATCAATTGTAG 8807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 133 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 192
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8808 ACTCAACGAGACGAAACCAAGAGAGTCTTGTCCGAAAGAAATGTAATTTCTGTGAAAA 8867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 193 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8868 TACAATTTCAAGGAAACAAATCATCTTGAGATTAAGCCGAATTGTTGAAACCGG 8927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 253 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8928 AAGCTAAATGTGATGCTTTGAACTGATGTAATCTCAAGAGAAAAAGTGTGG 8987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 313 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 8988 CTATGCTGAATGTGATTGAAATGTGATGTAATTAATGATTAAGAAAAATGTAGA 9047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 373 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9048 AGCCAAATCGAAGACATGCTATTGTAGAACTACTGAAAACAAGAAATATCAAGCAAT 9107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 433 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9108 TACAAGAGCAAAAGTAATTTCTTGAAAAAGAGCGCTGCAAGCTGCAATATTC 9167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 493 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9168 GAATGTACATTTCCATGAGATGATATTAATAATTTGCAAGACAGATGCAATCTTC 9227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 553 CCCCCCNRNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 612
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9228 TCAATTAATTAACCAAAATTTCAAAAAATATATTACTTTAGCATCAAGTAGTGAATCGAA 9287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 613 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9288 TGTGAATGTTCAATGGGTGCTGAAGCAGAGGTGCGAGAGCAACGACGATTGTGT 9347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 673 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 732
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9348 AAAAGATCAATTAATAGAACTAATATGTTGATTTAGCGAGAAAAATTAACAAGAGG 9407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 733 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 792
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9408 AAACATCAATGTATATGCCGATATGATTAATAATTAATTAAGTAGCAATTTCTAA 9467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 793 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 852
QY 9468 GCGCTATTGCGATGCGCAAAAGTCATGCTGAGCTGCTTCCGGAACCTGCACATTGAAA 9527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 853 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 912
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9528 AAATGAAGTAAATTTAATGATCGATCCGAGAAATTTAAATAATCTGCAAGATTGGA 9587
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 913 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 972
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9588 AGGAAAGCTAATAAAAAACCTCCGTGAGATCTAATCAGTAGCTGTATACGATTA 9647
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 973 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9648 ATATACATGCGCATCTTCTCGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGG 9707
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1032 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 9708 AGAAAAAGGAAAAAA 9723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1092 AAAAAAAAAAAAAAAAAAAAA 1107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 34  
DR147304  
LOCUS  
DEFINITION  
49027161 Drosophila pseudoobscura embryonic cDNA library Drosophila  
pseudoobscura cDNA clone O11 3', mRNA sequence.  
DR147304  
VERSION  
DR147304.1 GI:67892394  
KEYWORDS  
EST.  
ORGANISM  
Drosophila pseudoobscura  
Drosophila pseudoobscura  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 2071)

## REFERENCE AUTHORS

Richards, S., Liu, Y., Bettencourt, B.R., Hradecky, P., Letovsky, S.,  
Nielsen, R., Thornton, K., Hubisz, M.J., Chen, R., Weisell, R.P.,  
Coutome, O., Hua, S., Smith, M.A., Zhang, P., Liu, J., Buesemeyer, H.J.,  
van Batenburg, M.F., Howells, S.L., Scherer, S.E., Soderstrom, B.,  
Matthews, B.B., Crosby, M.A., Schroeder, A.J., Ortiz-Barrientos, D.,  
Rives, C.M., Metzker, M.L., Muzny, D.M., Scott, G., Steffen, D.,  
Wheeler, D.A., Worley, K.C., Hawjak, P., Dubchak, K.J., Egan, A.,  
Gill, R., Hume, J., Morgan, M.B., Miller, G., Hamilton, C., Huang, Y.,  
Malden, L., Verdusco, D., Clerc, Blenkinsburg, K.P., Dubchak, I.,  
Noor, M.A., Anderson, M., White, K.P., Clark, A.G., Schaeffer, S.W.,  
Gelbart, W., Weinstein, G.M., and Gibbs, R.A.  
Comparative genome sequencing of *Drosophila pseudoobscura*:  
chromosomal, gene, and cis-element evolution  
Genome Res. 15 (1), 1-18 (2005)

## JOURNAL PUBMED COMMENT

Contact: Stephen Richards  
Human Genome Sequencing Center  
Baylor College of Medicine  
One Baylor Plaza, Houston, TX 77030, USA  
Tel: 713-798-6667  
Email: stephenr@bcm.tmc.edu  
NCBI Trace Archive: 226691688  
Insert length: 1750 Std Error: 0.25.  
Location/Qualifiers

## FEATURES Source

1. 2071  
/organism="Drosophila pseudoobscura"  
/mol\_type="mRNA"  
/db\_xref="taxon:7237"  
/clone="O11"  
/dev\_stage="0-18h embryo"  
/clone\_1lb="Drosophila pseudoobscura embryonic cDNA  
library"  
/note="Vector: pOTB7; Site\_1: EcoRI; Site\_2: XhoI; oligo  
dt priming from poly A+ RNA, directionally cloned"

## ORIGIN

Query Match 0.9%; Score 87.8; DB 8; Length 2071;  
 Best Local Similarity 41.0%; Pred. No. 1e-06;  
 Matches 707; Conservative 0; Mismatches 1016; Indels 2; Gaps 1;

QY 7110 AGAAGATGTAATGTAAACGCAAAAGAGAAAAATATTACCAACACGACCAATATGC 7169  
 DB 252 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 311  
 QY 7170 AGAATCGAGAGCGACATCGAGCCAAATGCTTGTATTAATTTTGGAAACAGCTGT 7229  
 DB 312 AAAAAAGAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 371  
 QY 7230 AGAAGATGAAAAATTTCTGAGAAAAAGACAGAAATTTTAAACCTTAGACGAAT 7289  
 DB 372 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 431  
 QY 7290 TAAACAAGAACAGATTAATAATGATGCTACGAAAAATCTTAATCAGACAG 7349  
 DB 432 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 491  
 QY 7350 TATTTCTACAGAGATCTTCTGTAAGCGATAGAGAGATCTCAGGAGAGAGAT 7409  
 DB 492 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 551  
 QY 7410 TAAAGCATTTGAGAGCTTCTGATATTATTGAAAAATGTAGATTATTAACAGAG 7469  
 DB 552 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 611  
 QY 7470 CAGAAATTAATATCTTCTACTGCTGTTGGAGACTGACAGCTTCTGCTCCGATCAG 7529  
 DB 612 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 671  
 QY 7530 AACAGTGGCATTAACAATTTAAAGAAATTCGGAGTTACTGTTGAAATTTCTTTGT 7589  
 DB 672 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 731  
 QY 7590 GAAAGAGCTGAAAAATGTAATGTTAGTCGATTTTACGAAATGTTGCTTAAACG 7649  
 DB 732 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 791  
 QY 7650 ATATCAAGGCTCTGTAGAGCATTTGGAAATAGAGCTGCTATGACAAATTAATTTCTAA 7709  
 DB 792 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 851  
 QY 7710 TGGAGATCAAAATATCAATTAATAATTTCTAACTATTAGAAAAATATTGATGTAT 7769  
 DB 852 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 911  
 QY 7770 TGTAAAGATTAATTCGAAATTTGAGCGCAAGCAAAAGATTAACTGTAGACGGTAGC 7829  
 DB 912 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 971  
 QY 7830 TGGCGAGCATTAATCTCAAAAGCAAAAGATGAATTTAGAGTTGAATTTAGAA 7889  
 DB 972 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1031  
 QY 7890 GAGTATTTTCAATGAAGAAATAGAGTAAGTACCTCTTAAGGAATTTGGAAGAAAT 7949  
 DB 1032 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1091  
 QY 7950 CAATGTCAAAGTGGAAAAAGAAACAGAGTACTGCTGAATCTCAAGAGCTTCTGTAG 8009  
 DB 1092 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1151  
 QY 8010 AGCAGTAGCAGGCGAGAAATTTTCCGAACAAAGATGCGGAAGCTTATTTTGA 8069  
 DB 1152 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1211  
 QY 8070 AGTTAGTCAAAATCCGGAAGATTTTTCATGCAATATGTGATATGGAACCAAC 8129  
 DB 1212 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1271  
 QY 8130 ACATTAATGAAGTAACGCAATTTCTAAAGCATTAACAGTTCTGTATTTGGAGAGT 8189

DB 1272 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1331  
 QY 8190 TGAAGTACCAAGCGAAGCTATGCTGACGATAACTATGATGAAGTTGAGAGAG 8249  
 DB 1332 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1391  
 QY 8250 AATTTGTTCAAGCAAAATCGATTGAATCTTAAGTGAAGTTTGGATGAGAG 8309  
 DB 1392 GAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1451  
 QY 8310 TAAAGTAACTGTTAAATCTTCTGTAGTATCAGAAAAATGAGAGAAATGCGAGCAG 8369  
 DB 1452 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1511  
 QY 8370 AGTAATCTTCTACAGCAAAAGTAAATCTGAATCTGATTTCTTTACGAAGCAG 8429  
 DB 1512 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1571  
 QY 8430 TTAATGAATTAATGATTAACAAAAATATATTTCAGAACTAATGCTCTTGTATA 8489  
 DB 1572 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1631  
 QY 8490 TGATCAAAAGATGAAGCGAATATAGAAATCTTAAGCGTAGCCGTGTCATGACAG 8549  
 DB 1632 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1691  
 QY 8550 AACAAACAAAGATTTACAGAGATCAACAAAGTTAACTTCTACACTGTAATGAGAGAA 8609  
 DB 1692 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1749  
 QY 8610 CGTATCTCACTTCTGCAAAAGCTTTGCTTAAATGAAATTTATGAAATGTAAAG 8669  
 DB 1750 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1809  
 QY 8670 AACTGAGAGCCTTGTGCGAGCGGAAACAGAGCCGTTGAAATTTATACAAAGATAC 8729  
 DB 1810 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1869  
 QY 8730 TACAGAGCATTTGTTGCAAGAAATTTGGAAATTTGAGATTAATTAAGAACTTGCAG 8789  
 DB 1870 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1929  
 QY 8790 AGTAAATACATTTGTAAGTCAACAGAGACGAACCAAGAGAG 8834  
 DB 1930 AAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1974

RESULT 35  
 CL082658 1457 bp DNA linear GSS 31-DEC-2003  
 LOCUS CH216-169B13.RM4.1 CH216 Xenopus tropicalis genomic clone  
 DEFINITION CH216-169B13 genomic survey sequence.  
 ACCESSION CL082658  
 VERSION CL082658.1 GI:40538571  
 KEYWORDS GSS.  
 SOURCE Xenopus tropicalis (western clawed frog)  
 ORGANISM Xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 Xenopodinae; Xenopus; Silurana.  
 REFERENCE 1 (bases 1 to 1457)  
 AUTHORS Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,  
 Mardis, E. and Wilson, R.  
 TITLE A physical map of the xenopus tropicalis genome  
 JOURNAL Unpublished (2003)  
 COMMENT Contact: Richard K Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submls@wustl.edu  
 Insert Length: 17500 Std Error: 0.00  
 Seq primer: RM4 ctcaagcgatcggtcgcgagc  
 Class: BAC ends

High quality sequence start: 393  
High quality sequence stop: 1347.  
Location/Qualifiers

## FEATURES

source 1..1457  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/strain="Nigerian frog"  
/db\_xref="taxon:8364"  
/clone="CH216-169B13"  
/sex="male"  
/cell\_line="Stock 248 F7A2, indred N7"  
/clone\_lib="CH216"  
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

## ORIGIN

Query Match 0.94; Score 87.6; DB 10; Length 1457;  
Best Local Similarity 40.04; Pred. No. 1.1e-06;  
Matches 516; Conservative 0; Mismatches 773; Indels 0; Gaps 0;

```

8435 AAAATATGATTACACAAAAAATATATTTGAGAACTCATGCTCTTGCTTAAATGATA 8494
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
75  AAAAAAAAAAAGACCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 134
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8495 CAAAGATGAGGAAATATGAACTTTAGCGGTAGCCGGTGTGATGACAGAGACAA 8554
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135  NCCCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 194
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8555 ACAAGACATTTACGATACAAAGTTAACTTCTCAACTGTAAATGAGAGAAAGTAT 8614
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8615 CTCAACTGCTGCAAAAGCTTTGCTAAATAATGAAATTTGAAATGTAAGAAAGACTG 8674
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
255  AGAAAAAGAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8675 GAGGAGCCTTAGTCGAGCGGAAACGACGCCGTTGAATAATATACAAAGACTACTACG 8734
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
315  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 374
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8735 GAGCATTTGTTGAGAAATTTGGAAATTTGGAGATTAATTAAGAAAGATTGCAAGATA 8794
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
375  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 434
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8795 ATACGATTTGTAAGATCAACGAGACGAAACCAAGAGAGTCTTTCGGAAGAAATGTA 8854
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
435  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 494
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8855 TTTCTGTGAAAAATACATTTTCAGGGAAACAAATCATCTGATGAAGTAAAGCCAGA 8914
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
495  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 554
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8915 TTGTTGGAACCGGAATGTAAATGTAGATGCTTTGATGAATCACTGTATGATCTACAG 8974
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
555  GAAAGGAAAAAAGAAAAAAAAAAAAAAAAAGAGAGAAAAAAGAAAAAAAAAAAAAAAAA 614
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
8975 GAAAGAGTGTGCTATGTTGTAATTTGTAATTTGTAATTTGTAATTTGTAATTTGTAAT 9034
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
615  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 674
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
9035 AGAAAAATGTAGAGCAAAATCGAAGACATCTATTGTAGAACTACTGAAAAACAAG 9094
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
675  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 734
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
9095 AATATCAAGATTTACAGAGCAAAAGTAAATTTCTTGAAGAAAGAGACGCTGCAGCTG 9154
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
735  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 794
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
9155 CAGCTCAATATGAAATGTACATTTCCATGATGATGATATTAATAATTTGGCAAGC 9214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
795  AAAAAAAAAAAAAAAAAAGAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 854
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
9215 AGATGATCTTCTCAATTAATTAACCAAAATTCAAAAAATATATTAATTAATTAATTA 9274
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 855 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 914
Qy 9275 GTAGTGAATCGAATGTGATGTTTCATGGGTCGCTGAAAGCAAGAGATGACGCAAG 9334
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 915 AAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 974
Qy 9335 CGACAGATTGTTGTAAGATCAATTAATTAAGAACTTAATTAATTTGATTAGCAGAAAAA 9394
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 975 AAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1034
Qy 9395 TTAATAACAGAGGAAACATCAATGTATATGCCGATATGATTAATAATTAATTAAGTA 9454
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1035 AAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAA 1094
Qy 9455 AGCAAAATTTCTAAGGCTATTGGCGATGCGAAAGTCATGCTGACGCTTGCGCAACTG 9514
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1095 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1154
Qy 9515 CCACATTTGAAAAAATGAATTAATTAATTAATGCGATCCGAAATTTAAATAATTC 9574
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1155 AAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1214
Qy 9575 TGGCAAGATTGGAAGGAAAGCTAATTAATAAAAGTCCGTAGATCTAATCAGTAGACT 9634
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1215 AAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAA 1274
Qy 9635 GGTATACGATTAATATATCATGATGCTCTTCTGAAATAAGCATACAAAAATTTGACATATC 9694
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1275 AAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAA 1334
Qy 9695 AATCAAGAGAGAGAAAAAGGAAAAA 9723
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1335 AAAAAAAAAAGAAAAAAAAAGAAAAA 1363
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 36  
LOCUS CL081978 1428 bp DNA linear GSS 31-DEC-2003  
DEFINITION CH216-165F21\_Sp5.1 CH216 Xenopus tropicalis genomic clone  
ACCESSION CL081978  
VERSION CL081978  
KEYWORDS CL081978.1 GI:40537891  
SOURCE GSS.  
ORGANISM Xenopus tropicalis (western clawed frog)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
Xenopodinae; Xenopus; Silurana.  
REFERENCE 1 (bases 1 to 1428)  
Kremliński, C., Carter, J., McPherson, J., Warren, W., Graves, T.,  
Mardis, E. and Wilson, R.  
A physical map of the xenopus tropicalis genome  
Unpublished (2003)  
CONTACT: Richard K Wilson  
Genome Sequencing Center  
Washington University School of Medicine  
Email: submisio@wustl.edu  
Insert Length: 175000 Std Error: 0.00  
Seq primer: SPS atcgcgcgttcgaccc  
Clas: BAC ends  
High quality sequence start: 309  
High quality sequence stop: 527.  
Location/Qualifiers

## FEATURES

source 1..1428  
/organism="Xenopus tropicalis"  
/mol\_type="genomic DNA"  
/strain="Nigerian frog"  
/db\_xref="taxon:8364"  
/clone="CH216-165F21"  
/sex="male"  
/cell\_line="Stock 248 F7A2, indred N7"  
/clone\_lib="CH216"

Query Match	0.9%	Score 87.4	DB 10	Length 1428
Beet Local Similarity	37.5%	Pred. No. 1.2e-06		
Matches 496	Conservative	0	Mismatches 826	Indels 0
			Gaps 0	
QY 8402	AATCGTGTGTTGCTTTACGAAAGCAAGATTTATGAAATTAATGATTATACCAAAAAATATA	8461		
DB 96	AA	155		
OY 8462	TTTCGAGAGTCAATCTCTTGCTTTAATGATACAAAGAAATGAAGCAATATATGATCTT	8521		
DB 156	AA	215		
OY 8522	TAGCGGTAGCCGGTGTGATGCACAGAGACAAACAAAGCATTTATACGATCAACAGAT	8581		
DB 216	AA	275		
OY 8582	TAACTTCTACACATGTAATGAGAGAAAGTATCTCAACTTCGTCAGAAAGCTTTGGCTA	8641		
DB 276	AA	335		
OY 8642	AAATGAAATTAATGAGAAATGTAAGAGAACTGAGAGAGCTTATGCGAGCGGAAACAG	8701		
DB 336	AA	395		
OY 8702	CAGCGGTGAATAATATATACAGAGAGTACTACAGAGCATGTGGTTGCAGAAATTTGGGAA	8761		
DB 396	AA	455		
OY 8762	TTGAGATTAATTAGAAAGATTCAGAGATTAATACGATTGTAGAGTCAACGAGACG	8821		
DB 456	AA	515		
OY 8822	GACCAAGAGAGCTTGTGCGAAAGATGTGTATTTCTGTGAAAAATTCATTTCAAGGCG	8881		
DB 516	AA	575		
OY 8882	AAACAAATCATCTCATGTAAGTAAAGCCAGATTTGTAACCGGAATGTAAATGTAG	8941		
DB 576	AA	635		
OY 8942	ATGCTTGAATGAACTTGATGTAGATCTACAGAAAGAAAGTGCTATGTGGAATTG	9001		
DB 636	AA	695		
OY 9002	GTAATTGCAATGTTGATGTAAATATGATTAAGAAATATGTAAGCCAAATTCGGA	9061		
DB 696	AA	755		
OY 9062	GACATGCTATGTGAAACTACTGGAACCAAGATATCAAGCATTTACAGAGCAAAAG	9121		
DB 756	AA	815		
OY 9122	TAAATATTCTGAAAAAGAGACGCTGCAGCTGCAGTATCGAATGTACATTCATTT	9181		
DB 816	AA	875		
OY 9182	COATGAGATGATTAATAAATTTGGCAAGCATATGATCTTCTCAATTAATAACA	9241		
DB 876	AA	935		
OY 9242	AAATTTCAAAAAATTAATTAATCTTATGATCAATGATGATTCGAATGTGATGTCATG	9301		
DB 936	AA	995		
OY 9302	GGGTGCTGAAGCAGAGTGCTCAGAGCCAAAGCAGCTTATGTGTAAGATCAATTA	9361		
DB 996	AA	1055		
OY 9362	ATGAGACTAATTAATGTTGATTTAGCAGAAAAATTAACAGAGGAAACATCAATGAT	9421		

Dd	1056	AAA	1115
Oy	9422	ATGCCGATATGATAAAATTATATATTAAGTAGACAAATTCTAAGCGTATTCGGATG	9481
Dd	1116	AANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	1175
Oy	9482	CCAAAAGTCATGCTGCAGCTGTTCCGCCACTGCCACTATTTGAIAAAAAAAAAAGAACTAAAT	9541
Dd	1176	AAA	1235
Oy	9542	TTAATTAATGCGATCCGGAATTTAAATAATCTGCGAAGTTGAAAGCGAAGCTAATA	9601
Dd	1236	AAA	1295
Oy	9602	AAAAAAAAAGCGGTAGATCTAATAGGTAGA CTGTATACGATTAATATATCATGCAATT	9661
Dd	1296	AAA	1355
Oy	9662	CTTCTGAAAAAGCATTACAAAAAATTGACATATCAATCAAAGAGAGGAAAAAGGAAAA	9721
Dd	1356	AANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA	1415
Oy	9722	AA 9723	
Dd	1416	AA 1417	

RESULT 37					
AJ927291/c					
LOCUS	AJ927291	1139 bp	mRNA	linear	EST 10-JUN-2005
DEFINITION	AJ927291	<i>Theileria annulata</i> merozoite <i>Theileria annulata</i> cDNA clone.			
	ta026d02_g1k,	mRNA sequence.			
ACCESSION	AJ927291				
VERSION	AJ927291.1	GI:67497675			
KEYWORDS	EST.				
SOURCE	<i>Theileria annulata</i>				
ORGANISM	<i>Theileria annulata</i>				

REFERENCE	1 (bases 1 to 1139)
AUTHORS	Pain, A., Renaud, H., Berriman, M., Murphy, L., Yeats, C.A., Weir, W.

Kermottou, A., Aslett, M., Bishop, R., Bouchner, C., Coche, M.,  
 Coulson, R.M.R., Cronin, A., de Villiers, E., Fraser, A., Foster, N.,  
 Gardner, M., Goble, A., Griffiths-Jones, S., Harris, D.E., Katzer, F.,  
 Larke, N., Lord, A., Maser, P., McCellar, S., Mooney, P., Morton, F.,  
 Nene, V., O'Neill, S., Price, C., Quail, M.A., Rabinowitsch, E.,  
 Rawlings, N.D., Rutter, S., Saunders, D., Seeger, K., Shah, T.,  
 Squares, R., Squares, S., Tivey, A., Walker, A.R., Woodward, J.,  
 Dobbelaere, D.A.E., Langsley, G., Rajandream, M.-A., McKeever, D.,  
 Shleis, B., Tait, A., Barrall, B. and Hall, N.  
 The genome of the host-cell transforming parasite *Theileria*  
*annulata* and a comparison with *T. parva*  
 Unpublished (2005)

**FEATURES**  
The Pathogen Sequencing Unit  
The Wellcome Trust Sanger Institute  
Genome Campus, Cr10 1SA, UNITED KINGDOM  
Microicide cDNA library: Frank Katzer, Division of  
Veterinary Infection and Immunity, ICM, University of Glasgow, UK.  
Location/Qualifiers  
1. .1139

```

/organism="Theileria annulata"
/mol_type="mRNA"
/isolate="Ankara (clone D7)"
/db_xref="taxon:5874"
/clone="tam028d02.g1k"
/dev_stage="merozoite"
/lab_host="Bos taurus (cow)"
/clone_id="Theileria annulata merozoite"
/note="Country: Turkey:Ankara"

```

**ORIGIN**







```

Db      1058 AAAAAAAAAAAAAAAAAACAAAAAAAAAAAAAAAAACAGAGAAAAACA 1117
QY      9381 TTAGACAGAAAAATTAAAAACAGAGGAAACATCAATGTATATGCCGATATGATAAAA 9440
Db      1118 AAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAACAAAAACAAAAACA 1177
QY      9441 TTATATATATAGTACAAATTCTAAGGCTATTGCCGATGCCAAAAGTCATGCTGCAGC 9500
Db      1178 AGAACAAAAAAGAAAAAGAACAAACAAAAAAGAAAAAAGAAAAAAGA 1237
QY      9501 TGCTTCGCAACTGCCACTATTGAAAAAATGAAGTAAATTTAATATGCGATCCGAGA 9560
Db      1238 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACA 1297
QY      9561 ATTTAAAAATATCTGCCAAGATTGGAAGGAAAGCTAATTAATAAAACGTCCGTAGGATC 9620
Db      1298 AAAAAAAAAAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 1357
QY      9621 TAATCAGGTAGACTGGTATACGATTAATATACATGCGATTCTTGAAAAAGCATACAA 9680
Db      1358 AAAAGACAAAAAACAACAAAAACAAGAAAAAAGAAAAAAGAAAAATATACAAAG 1417
QY      9681 AAAATTGACATATCAATCAAGAGAGAAAAAGGAAAAAA 9723
Db      1418 AAAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1460

```

Search completed: November 25, 2005, 20:31:17  
 Job time : 22535 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2005, 14:17:17 ; Search time 4501 Seconds  
(without alignments)  
17868.911 Million cell updates/sec

Title: US-10-647-057-8

Perfect score: 9726  
Sequence: 1 atgagcgagcacaataa.....gagaaaggaataatga 9726

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications NA.Main:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	9726	100.0	9726	3	US-09-841-786-8 Sequence 8, Appl1
2	9726	100.0	9726	3	US-10-647-057-8 Sequence 8, Appl1
3	9726	100.0	11130	7	US-09-841-786-15 Sequence 15, Appl1
4	9726	100.0	11130	7	US-10-647-057-15 Sequence 15, Appl1
5	2780	28.6	2780	3	US-09-841-786-10 Sequence 10, Appl1
6	2780	28.6	2780	7	US-10-647-057-10 Sequence 10, Appl1
7	2322	23.9	2322	3	US-09-841-786-13 Sequence 13, Appl1
8	2322	23.9	2322	7	US-10-647-057-13 Sequence 13, Appl1
9	2141	22.0	2141	7	US-09-841-786-11 Sequence 11, Appl1
10	2141	22.0	2141	7	US-10-647-057-11 Sequence 11, Appl1
11	1887	19.4	1887	3	US-09-841-786-12 Sequence 12, Appl1
12	1887	19.4	1887	7	US-10-647-057-12 Sequence 12, Appl1
13	1130	11.6	1130	3	US-09-841-786-9 Sequence 9, Appl1
14	1130	11.6	1130	7	US-10-647-057-9 Sequence 9, Appl1
15	116.2	1.2	5152	3	US-09-742-096-2 Sequence 1, Appl1
16	116.2	1.2	5152	7	US-10-415-253-1 Sequence 1, Appl1
17	107.6	1.1	5528	7	US-10-668-749A-1 Sequence 1, Appl1
18	80.4	0.8	1300	7	US-09-815-242-4761 Sequence 4761, Ap
19	79.2	0.8	7434	8	US-10-470-048B-201 Sequence 201, Ap
20	79.2	0.8	7437	3	US-09-815-242-8869 Sequence 8869, Ap
21	79.2	0.8	8155	2	US-08-781-986A-63 Sequence 63, Appl
22	79.2	0.8	8155	7	US-10-329-624-63 Sequence 63, Appl
23	79.2	0.8	8155	7	US-10-329-624-63 Sequence 63, Appl

24	74.8	0.8	7104	3	US-09-815-242-4580 Sequence 4580, Ap
25	74.8	0.8	7107	3	US-09-815-242-8291 Sequence 8291, Ap
26	72.2	0.8	600	9	US-10-972-079-13288 Sequence 13288, A
27	72.4	0.7	7446	7	US-10-282-122A-7578 Sequence 7578, Ap
28	70.6	0.7	11130	8	US-10-425-115-132399 Sequence 132399, A
29	68.8	0.7	31124	5	US-10-087-192-463 Sequence 463, App
30	68.2	0.7	4911	8	US-10-470-048B-427 Sequence 427, App
31	68.2	0.7	4997	7	US-10-282-122A-35506 Sequence 35506, A
32	67.6	0.7	7442	7	US-10-221-714A-409 Sequence 409, App
33	64.8	0.7	1891	3	US-09-742-096-4 Sequence 4, Appl1
34	64.4	0.7	1712	7	US-10-415-253-3 Sequence 3, Appl1
35	63.8	0.7	1121	8	US-10-425-115-115706 Sequence 115706, A
36	63.8	0.6	996	8	US-10-425-115-151933 Sequence 151933, A
37	62.6	0.6	673	7	US-10-424-599-57885 Sequence 57885, A
38	62.4	0.6	1062	8	US-10-425-115-120013 Sequence 120013, A
39	62.4	0.6	1243	8	US-10-706-635-24 Sequence 24, Appl
40	62.4	0.6	50000	7	US-10-021-323-9336 Sequence 9336, Ap
41	62.2	0.6	627	2	US-10-450-763-11257 Sequence 11257, A
42	61.6	0.6	441	9	US-10-311-455-1670 Sequence 1670, Ap
43	61.4	0.6	6668	6	US-10-311-455-2231 Sequence 2231, Ap
44	61.2	0.6	12237	6	US-10-311-455-2231 Sequence 2231, Ap
45	61	0.6	626	8	US-10-357-930-60960 Sequence 60960, A

## ALIGNMENTS

```
RESULT 1
US-09-841-786-8
; Sequence 8, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 9726
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-8

Query Match      100.0%; Score 9726; DB 3; Length 9726;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGAGCGGCGATCAAAATTAACGTTCAAGAGACAAAGAAAGATACAGATCTTAATAAAA 60
        1 ATGAGCGGCGATCAAAATTAACGTTCAAGAGACAAAGAAAGATACAGATCTTAATAAAA 60
        1 ATGAGCGGCGATCAAAATTAACGTTCAAGAGACAAAGAAAGATACAGATCTTAATAAAA 60
QY      61 GTTTTAAATGATTTGGGATTTGTTGATTAACATATGACGTTAGGCTATGATCAATC 120
        61 GTTTTAAATGATTTGGGATTTGTTGATTAACATATGACGTTAGGCTATGATCAATC 120
        61 GTTTTAAATGATTTGGGATTTGTTGATTAACATATGACGTTAGGCTATGATCAATC 120
Db      61 GTTTTAAATGATTTGGGATTTGTTGATTAACATATGACGTTAGGCTATGATCAATC 120
QY      121 ACCGCGACTGAGAAATTTGGAACAAATAATAGAAAAAAGAGATATGTTATGACATTTACT 180
        121 ACCGCGACTGAGAAATTTGGAACAAATAATAGAAAAAAGAGATATGTTATGACATTTACT 180
        121 ACCGCGACTGAGAAATTTGGAACAAATAATAGAAAAAAGAGATATGTTATGACATTTACT 180
Db      121 ACCGCGACTGAGAAATTTGGAACAAATAATAGAAAAAAGAGATATGTTATGACATTTACT 180
QY      181 ACAAACAAATTCAGGAGGAGAAACGTTTAAACAGTTTAAATGATTTGCTTTAACAAGAA 240
        181 ACAAACAAATTCAGGAGGAGAAACGTTTAAACAGTTTAAATGATTTGCTTTAACAAGAA 240
        181 ACAAACAAATTCAGGAGGAGAAACGTTTAAACAGTTTAAATGATTTGCTTTAACAAGAA 240
Db      181 ACAAACAAATTCAGGAGGAGAAACGTTTAAACAGTTTAAATGATTTGCTTTAACAAGAA 240
QY      241 AATAATATGCAAAATCATATATTTGGGAAAAAGATAGTACGGGGCTAATATATCTTTT 300
        241 AATAATATGCAAAATCATATATTTGGGAAAAAGATAGTACGGGGCTAATATATCTTTT 300
        241 AATAATATGCAAAATCATATATTTGGGAAAAAGATAGTACGGGGCTAATATATCTTTT 300
```

Db 241 AATAATATACAAATCTATATTTGGGAGAAAGAAATAGTACGGGGTAATAATCTTTT 300  
Qy 301 AACTTGTCAATGAAAAATTTGAAGTAGATGGATTATCAACGGAATTCGAGAAATATAA 360  
Db 301 AACTTGTCAATGAAAAATTTGAAGTAGATGGATTATCAACGGAATTCGAGAAATATAA 360  
Qy 361 ATTGGAGGAATTTATATTTCTTAAAGCTCGAAGGGATGCGATGAGAAAAATGAGATT 420  
Db 361 ATTGGAGGAATTTATATTTCTTAAAGCTCGAAGGGATGCGATGAGAAAAATGAGATT 420  
Qy 421 ATCAATGCGGTCTTTTCATCTATTTATCCAAAACAAGATTTTAAAGAGCTTTG 480  
Db 421 ATCAATGCGGTCTTTTCATCTATTTATCCAAAACAAGATTTTAAAGAGCTTTG 480  
Qy 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGGAATCATTCAGATGATGAAAGATATAA 540  
Db 481 GAAGAAGCCAAACATGTTAAAGTTTAAATGGAATCATTCAGATGATGAAAGATATAA 540  
Qy 541 ATTCCATGGAATCCGAATGGAAGCATTAACGTTAGAAAGAAAAATCAATGCTGTGAAGC 600  
Db 541 ATTCCATGGAATCCGAATGGAAGCATTAACGTTAGAAAGAAAAATCAATGCTGTGAAGC 600  
Qy 601 ATCGTTTATATGCGCGCGGATATTAGATTGAAGATCTGCAATCTAAGACAGAAATT 660  
Db 601 ATCGTTTATATGCGCGCGGATATTAGATTGAAGATCTGCAATCTAAGACAGAAATT 660  
Qy 661 ACAGATTTTAAATTTAGTCAATATTAGTATGATCGAATTAATTTGCTGACCGGAAT 720  
Db 661 ACAGATTTTAAATTTAGTCAATATTAGTATGATCGAATTAATTTGCTGACCGGAAT 720  
Qy 721 TTTAAAGCTACCAAGACAAATCTGAGATATTATTTCTTACGCTCAATAGATTCTCT 780  
Db 721 TTTAAAGCTACCAAGACAAATCTGAGATATTATTTCTTACGCTCAATAGATTCTCT 780  
Qy 781 CAAAAGCTATGCGGAAAAATTTCACTGTGTAAGAAAGAAATGAAAGATATTGTAAGAA 840  
Db 781 CAAAAGCTATGCGGAAAAATTTCACTGTGTAAGAAAGAAATGAAAGATATTGTAAGAA 840  
Qy 841 AATACCAAGCAAAATATTGAATCTGATCTGATTTGAAGCAGATGAAATATAAAATT 900  
Db 841 AATACCAAGCAAAATATTGAATCTGATCTGATTTGAAGCAGATGAAATATAAAATT 900  
Qy 901 AGTGGAAAGCTACAAATGCGAGATTTTAAAGAAAGAGGGGAAAAAGAACTTATAC 960  
Db 901 AGTGGAAAGCTACAAATGCGAGATTTTAAAGAAAGAGGGGAAAAAGAACTTATAC 960  
Qy 961 ACTCCTTTAAGTTTATGAGATGCTGAGATCTCCGTAAGATTAATTAAGAAAGTCTA 1020  
Db 961 ACTCCTTTAAGTTTATGAGATGCTGAGATCTCCGTAAGATTAATTAAGAAAGTCTA 1020  
Qy 1021 GGAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTTAGTACT 1080  
Db 1021 GGAAGAAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTTAGTACT 1080  
Qy 1081 AAGCTTGCAGAACACTTTTATGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAT 1140  
Db 1081 AAGCTTGCAGAACACTTTTATGCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAT 1140  
Qy 1141 GGAATTTTAAAGTTTATGCAAGATGCTGAGATCTGTTATTTGAAAGAAATGCCAAATC 1200  
Db 1141 GGAATTTTAAAGTTTATGCAAGATGCTGAGATCTGTTATTTGAAAGAAATGCCAAATC 1200  
Qy 1201 GAAGCAACAGAAAGAAAGCAATATTCATTTCTTACAGTGAAGTAAAGCACTATGGA 1260  
Db 1201 GAAGCAACAGAAAGAAAGCAATATTCATTTCTTACAGTGAAGTAAAGCACTATGGA 1260  
Qy 1261 GCAAGTACTTCTCATTTAAATTAACAATTTATATTGAGAAAGCCATGGAATCTT 1320  
Db 1261 GCAAGTACTTCTCATTTAAATTAACAATTTATATTGAGAAAGCCATGGAATCTT 1320  
Qy 1321 CTCAGTATTCGAGCGGATATATTTCTGCAAAAGATATTCGATGTAATTTGAAGA 1380  
Db 1321 CTCAGTATTCGAGCGGATATATTTCTGCAAAAGATATTCGATGTAATTTGAAGA 1380

Qy 1381 GAAGTAAATCGAAGGAGAGACAGATATTACTTCAAAATCTGAAAAATACTATTGATGCT 1440  
Db 1381 GAAGTAAATCGAAGGAGAGAGACAGATATTACTTCAAAATCTGAAAAATACTATTGATGCT 1440  
Qy 1441 TCTGTTCTGTTGGAACGATGAGATTTCAATTAAGTAGCTTTTCAGTATGCTGACG 1500  
Db 1441 TCTGTTCTGTTGGAACGATGAGATTTCAATTAAGTAGCTTTTCAGTATGCTGACG 1500  
Qy 1501 GAAGGAGAAATTAATCTTCGTCAGATTCGTAAGAGCAAAAGTATGAAATCGAAACG 1560  
Db 1501 GAAGGAGAAATTAATCTTCGTCAGATTCGTAAGAGCAAAAGTATGAAATCGAAACG 1560  
Qy 1561 GATGATGTAATGTTGAGAGTGAAGCCGATTATTCATTCGAGCTGCTGTAAAGTGA 1620  
Db 1561 GATGATGTAATGTTGAGAGTGAAGCCGATTATTCATTCGAGCTGCTGTAAAGTGA 1620  
Qy 1621 TTGGGGGATAGTGGTAAATGGGGTTGTGGCTGCAAAATATTTCTAATATATGCTTCTCC 1680  
Db 1621 TTGGGGGATAGTGGTAAATGGGGTTGTGGCTGCAAAATATTTCTAATATATGCTTCTCC 1680  
Qy 1681 CGTATAGATGTTAGATGATATCTACATGCCAAAGAGCACTAAATGTTGAGAGCTCATAC 1740  
Db 1681 CGTATAGATGTTAGATGATATCTACATGCCAAAGAGCACTAAATGTTGAGAGCTCATAC 1740  
Qy 1741 ATTACTAAAAATAGTGTCTGCAAAACAGATCTGATTTGGGAACCTCCAGTTTATGAT 1800  
Db 1741 ATTACTAAAAATAGTGTCTGCAAAACAGATCTGATTTGGGAACCTCCAGTTTATGAT 1800  
Qy 1801 GATCACTTTATGATCAAGTCAATCTAATCAATTTTATGATGCAATTAACACGCGTTT 1860  
Db 1801 GATCACTTTATGATCAAGTCAATCTAATCAATTTTATGATGCAATTAACACGCGTTT 1860  
Qy 1861 GGAAGAGACAGTGTCAATGAGAAATTAAGATTAAGCTTAAGAACTTAAAGTGTGCT 1920  
Db 1861 GGAAGAGACAGTGTCAATGAGAAATTAAGATTAAGCTTAAGAACTTAAAGTGTGCT 1920  
Qy 1921 GTGCTGCAACCATAGCAAAATCATATTAATCTGCTTCTGAGCAATAGAGAGATGGA 1980  
Db 1921 GTGCTGCAACCATAGCAAAATCATATTAATCTGCTTCTGAGCAATAGAGAGATGGA 1980  
Qy 1981 AGACTTTCTTCAAGAGTGAAGGAGATATGTAAGGCACTTAAATGAAGCTCAAAATCTT 2040  
Db 1981 AGACTTTCTTCAAGAGTGAAGGAGATATGTAAGGCACTTAAATGAAGCTCAAAATCTT 2040  
Qy 2041 CGAGCGACTACGTAAATGGAAGTGTGCTGTACGAAAGGAAAGAAAGAACTTAT 2100  
Db 2041 CGAGCGACTACGTAAATGGAAGTGTGCTGTACGAAAGGAAAGAAAGAAAGCTTAT 2100  
Qy 2101 GGAATGCAAGATTTTATGAGAACTATAAATATATGCTTCTGTCGACAAATGCGGAT 2160  
Db 2101 GGAATGCAAGATTTTATGAGAACTATAAATATATGCTTCTGTCGACAAATGCGGAT 2160  
Qy 2161 CATGCTGAATGTTGATGGAAGGAAATTTGATATCAACAGTGAATAAATATGAATAT 2220  
Db 2161 CATGCTGAATGTTGATGGAAGGAAATTTGATATCAACAGTGAATAAATATGAATAT 2220  
Qy 2221 AAAAATCTTCAAAAATGCGCAAAAGTGTGTTATTAATTAATTAAGAAAGAGCT 2280  
Db 2221 AAAAATCTTCAAAAATGCGCAAAAGTGTGTTATTAATTAATTAAGAAAGAGCT 2280  
Qy 2281 TTTGAAAAAGAAAGAAAGCTCCAGAAATATGATCCGAAAGATATGGAATCTATGGA 2340  
Db 2281 TTTGAAAAAGAAAGAAAGCTCCAGAAATATGATCCGAAAGATATGGAATCTATGGA 2340  
Qy 2341 TTATTTGAATGATTTTCAAGAAAAATTTGATGAAAAACCGAGCTTTTATCTAAATGTTGA 2400  
Db 2341 TTATTTGAATGATTTTCAAGAAAAATTTGATGAAAAACCGAGCTTTTATCTAAATGTTGA 2400  
Qy 2401 AGAATGACATTAATTTCTTCCGAGATGGAATTTCAAAAACAGAACTGCTATTAAGAAATTTGCA 2460  
Db 2401 AGAATGACATTAATTTCTTCCGAGATGGAATTTCAAAAACAGAACTGCTATTAAGAAATTTGCA 2460

Oy	2461	AACTATGCTCAGGAGAAATGAAAAAATTAAGAGAAAAATTCAGAAAGATTTAAAGCT	2520
Dp	2461	AACTATGCTCAGGAGAAATGAAAAAATTAAGAGAAAAATTCAGAAAGATTTAAAGCT	2520
Oy	2521	TTTTCAAGAGATTGACATGACCTGATTAAGAAACTTTGAATTTTACAGAGATGAGAAAT	2580
Dp	2521	TTTTCAAGAGATTGAGTGAATGATTAAGAAACTTTGAATTTTACAGAGATGAGAAAT	2580
Oy	2581	TATGCAAAATTTTACACTTTTACCTCTCCGAGCTAATGAGAAAGAGATGTTTCTTCT	2640
Dp	2581	TATGCAAAATTTTACACTTTTACCTCTCCGAGCTAATGAGAAAGAGATGTTTCTTCT	2640
Oy	2641	GTGGAGAGAGCTGTTTCGTGGGTAGAAACAGGAAATTATGCAAGATATCCGTTGGAAAA	2700
Dp	2641	GTGGAGAGAGCTGTTTCGTGGGTAGAAACAGGAAATTATGCAAGATATCCGTTGGAAAA	2700
Oy	2701	GGAGCTAAACTTGTGCAAAAAAGATTTAAATTAAGCTATCAATTAAGCAAAACA	2760
Dp	2701	GGAGCTAAACTTGTGCAAAAAAGATTTAAATTAAGCTATCAATTAAGCAAAACA	2760
Oy	2761	GTGAATTTAGTTGGAAATATTGCACTTGCGAAGACGATCATCCGAGTGCAGTCCGA	2820
Dp	2761	GTGAATTTAGTTGGAAATATTGCACTTGCGAAGACGATCATCCGAGTGCAGTCCGA	2820
Oy	2821	GGAAGATTAAATGTTCCAAAGATCGAAAAATTCAGCTATCGTAGAGCTAAAGAAAAAGCT	2880
Dp	2821	GGAAGATTAAATGTTCCAAAGATCGAAAAATTCAGCTATCGTAGAGCTAAAGAAAAAGCT	2880
Oy	2881	GAAATTTATCAGAGAAAAATTTAATGCAATGCAATTGAAACAGACTTTTCATGTAGCCGGA	2940
Dp	2881	GAAATTTATCAGAGAAAAATTTAATGCAATGCAATTGAAACAGACTTTTCATGTAGCCGGA	2940
Oy	2941	TCTTTTAAATGATGCTCAGTGGGAAATGCAATCAATGGAATGGAGTTTATGTAGAGGT	3000
Dp	2941	TCTTTTAAATGATGCTCAGTGGGAAATGCAATCAATGGAATGGAGTTTATGTAGAGGT	3000
Oy	3001	ATCAGTAAGGCAAGAGTTTCCATTGATGACGAGCATATTTGAAAGCTAATTAAAAAAAT	3060
Dp	3001	ATCAGTAAGGCAAGAGTTTCCATTGATGACGAGCATATTTGAAAGCTAATTAAAAAAAT	3060
Oy	3061	GCTTTTAAACAGTAAGATGATTACTTCTGTTTGGAAATGCTGCGGTTCAAGCCGGAATCCGA	3120
Dp	3061	GCTTTTAAACAGTAAGATGATTACTTCTGTTTGGAAATGCTGCGGTTCAAGCCGGAATCCGA	3120
Oy	3121	ACGAAAAATGCGGCGGTCGTGGGTTGCTGTCGCGTAATGATTTATTTCAACAA	3180
Dp	3121	ACGAAAAATGCGGCGGTCGTGGGTTGCTGTCGCGTAATGATTTATTTCAACAA	3180
Oy	3181	GCTTCATTTGAAGATTAATGACGAAGGACAAAGTAATATGATTAAGAAATTAAGATGATGAA	3240
Dp	3181	GCTTCATTTGAAGATTAATGACGAAGGACAAAGTAATATGATTAAGAAATTAAGATGATGAA	3240
Oy	3241	GTAACAGTAACGTCGCGAATCTTTAGAGTAGATGCAAAAACGACCGGAACATCAACAGT	3300
Dp	3241	GTAACAGTAACGTCGCGAATCTTTAGAGTAGATGCAAAAACGACCGGAACATCAACAGT	3300
Oy	3301	ATTTCTGTGTCGCGAGGAATTAATPAAGGTTGAAAGTAAACGAGTGAAGAAAAACGAA	3360
Dp	3301	ATTTCTGTGTCGCGAGGAATTAATPAAGGTTGAAAGTAAACGAGTGAAGAAAAACGAA	3360
Oy	3361	TCAGAAAGAAACCCAGAGGATTTTGTGCAAAATCGGAAACAAAGTGAATCTGTATAA	3420
Dp	3361	TCAGAAAGAAACCCAGAGGATTTTGTGCAAAATCGGAAACAAAGTGAATCTGTATAA	3420
Oy	3421	AATATAAATTAACGATATGATGATTCATTTAACGAAAAATTTCAATTTACATTTCTGAA	3480
Dp	3421	AATATAAATTAACGATATGATGATTCATTTAACGAAAAATTTCAATTTACATTTCTGAA	3480
Oy	3481	GGAGTAAGAAAAAGCGGGGAAATTTCTTCCGAAACGTTTCTCATCTCCCGATTAAGAGACG	3540
Dp	3481	GGAGTAAGAAAAAGCGGGGAAATTTCTTCCGAAACGTTTCTCATCTCCCGATTAAGAGACG	3540
Oy	3541	TCTTTCAAGTTGGAGCTTCTGGAAGTGTCTTTCAATATAATTAAGAAAAACAATCT	3600

Db	3541	TCCTTCAGTTTGGAGCTTCGAAAGTGTTCCTTCAATATATTTAAAAAGAAACATCT	3600
OY	3601	GCTGTCATAGATGAGTAAGATAATTAATTGAAAGGAGCAATAATAAGTAGAGTACT	3660
Db	3601	GCTGTCATAGATGAGTAAGATAATTAATTGAAAGGAGCAATAATAAGTAGAGTACT	3660
OY	3661	TCCTTCGATTCATCTTTTGTGTGAGACATGAGGGCGGATCTGCTGCATCTTCACTGGAATCAT	3720
Db	3661	TCCTTCGATTCATCTTTTGTGTGAGACATGAGGGCGGATCTGCTGCATCTTCACTGGAATCAT	3720
OY	3721	ATTGGAAAGTGAAATAGCAACATCAGTGCCTTTAGTCGAGAGCGCTGCCTTAATAAT	3780
Db	3721	ATTGGAAAGTGAAATAGCAACATCAGTGCCTTTAGTCGAGAGCGCTGCCTTAATAAT	3780
OY	3781	ATTCAAGGTAAAAACAAGCTTTGGTTAAAAATAGTATATTGCAATATGCAATTAATTT	3840
Db	3781	ATTCAAGGTAAAAACAAGCTTTGGTTAAAAATAGTATATTGCAATATGCAATTAATTT	3840
OY	3841	AAAGTAATATCTTTGATGAGAGAACTCAAGTACAGACAGAGAGAGCTTTGGAAGCATT	3900
Db	3841	AAAGTAATATCTTTGATGAGAGAACTCAAGTACAGACAGAGAGAGCTTTGGAAGCATT	3900
OY	3901	AAAGAAAGTGAGAGACAAAGAAAAAGTTATCTATTGGAACTTCTGCTTCTATCAACTTA	3960
Db	3901	AAAGAAAGTGAGAGACAAAGAAAAAGTTATCTATTGGAACTTCTGCTTCTATCAACTTA	3960
OY	3961	GTGAACAATGAAAGTTCTGCAAAAATCAAAAAATTAACAGTAGCAGAGAAATCGAAGC	4020
Db	3961	GTGAACAATGAAAGTTCTGCAAAAATCAAAAAATTAACAGTAGCAGAGAAATCGAAGC	4020
OY	4021	CAAAAAATGATGTTGATGTCACTGCTTATCAAGCGAGCACCCAAGTACAGAGACTTTA	4080
Db	4021	CAAAAAATGATGTTGATGTCACTGCTTATCAAGCGAGCACCCAAGTACAGAGACTTTA	4080
OY	4081	AATTTACAAGCTGAAAGTCAATATGAACTGTAGGGGCTACTGACTGTTGCCAAATTA	4140
Db	4081	AATTTACAAGCTGAAAGTCAATATGAACTGTAGGGGCTACTGACTGTTGCCAAATTA	4140
OY	4141	AACAACAAGTAATATGCTCTATATAGTGTGAGAGATATCTAACTTAAATCGACCGGAC	4200
Db	4141	AACAACAAGTAATATGCTCTATATAGTGTGAGAGATATCTAACTTAAATCGACCGGAC	4200
OY	4201	GCAAAAGCTCTTTTAGCAACCACTCAAGTACATGCTGCACTGACGAGCGGAGGACAAAT	4260
Db	4201	GCAAAAGCTCTTTTAGCAACCACTCAAGTACATGCTGCACTGACGAGCGGAGGACAAAT	4260
OY	4261	AGTTCTGAGCGGAGATTAGGAAATTAATCAAGGGGCTGTTCTGTCAATTAAGATTGACAT	4320
Db	4261	AGTTCTGAGCGGAGATTAGGAAATTAATCAAGGGGCTGTTCTGTCAATTAAGATTGACAT	4320
OY	4321	GACGTGGAAGCTACGTTGATTAATCTTCCATCGAAGAGCTTAATGAATTAATGTCAAT	4380
Db	4321	GACGTGGAAGCTACGTTGATTAATCTTCCATCGAAGAGCTTAATGAATTAATGTCAAT	4380
OY	4381	GCCAAAGATGCAAGGAAGTTCTGATCTAGCAAAAGAAATACAGGCTTTACTAAATGGA	4440
Db	4381	GCCAAAGATGCAAGGAAGTTCTGATCTAGCAAAAGAAATACAGGCTTTACTAAATGGA	4440
OY	4441	AAAGATAAAAAATTTTGAAGATCGTGTATTAAATCGACTGGAATAATGTTATTATACG	4500
Db	4441	AAAGATAAAAAATTTTGAAGATCGTGTATTAAATCGACTGGAATAATGTTATTATACG	4500
OY	4501	AAAGAAACAACATGAAAAAAGCAAAAGAAAAAGAAAGAGAGCGCTATTGTAAATGCTGCTTA	4560
Db	4501	AAAGAAACAACATGAAAAAAGCAAAAGAAAAAGAAAGAGAGCGCTATTGTAAATGCTGCTTA	4560
OY	4561	TCGGTTGCTGAAACGATTAATTCGGCTGGAAGATAGCTATTGCAAGTCAATACTGTTAAA	4620
Db	4561	TCGGTTGCTGAAACGATTAATTCGGCTGGAAGATAGCTATTGCAAGTCAATACTGTTAAA	4620
OY	4621	AATTAATTTAAGCAAAATTGATGTGAAGCAATAAGAAAGCCGAGAGGATTAATTCAT	4680
Db	4621	AATTAATTTAAGCAAAATTGATGTGAAGCAATAAGAAAGCCGAGAGGATTAATTCAT	4680

Db 4621 AATTAATTTAAAGCAATTGAGTGAGAACAAATAAGAAACCGAGAGATTAATTCAT 4680  
Qy 4681 GCGAAACATGTAAATGTGAGGCAAAATCATCTACTGTGTGTGAATGCGGCTTCTGGA 4740  
Db 4681 GCGAAACATGTAAATGTGAGGCAAAATCATCTACTGTGTGTGAATGCGGCTTCTGGA 4740  
Qy 4741 CTTCCTATCAGCAAAAGATGCTTTTTCAGGAATGGATCTGAGCATGCGAAGACTTTATCA 4800  
Db 4741 CTTCCTATCAGCAAAAGATGCTTTTTCAGGAATGGATCTGAGCATGCGAAGACTTTATCA 4800  
Qy 4801 AATGACACGATTCGCAAAAGTGATAAGGAAGAATTTCTGCTGATTTCTCTTAATGTTGAC 4860  
Db 4801 AATGACACGATTCGCAAAAGTGATAAGGAAGAATTTCTGCTGATTTCTCTTAATGTTGAC 4860  
Qy 4861 GCAAAATTAATTCATTTCTGAGGTGAATGTTGCGGGAACCATGCGGCTTCTCTTCTACG 4920  
Db 4861 GCAAAATTAATTCATTTCTGAGGTGAATGTTGCGGGAACCATGCGGCTTCTCTTCTACG 4920  
Qy 4921 GCGGTAGAGCTGCTTTTTCGCAATTAATCTTTCAATTAATAAACCCTGCTTTGATTACA 4980  
Db 4921 GCGGTAGAGCTGCTTTTTCGCAATTAATCTTTCAATTAATAAACCCTGCTTTGATTACA 4980  
Qy 4981 GGAAGCAAGCTAAATCTTTTGTAGTGAAGAAATACAAAGTCAATGTACAGCTTGTGAT 5040  
Db 4981 GGAAGCAAGCTAAATCTTTTGTAGTGAAGAAATACAAAGTCAATGTACAGCTTGTGAT 5040  
Qy 5041 GATTCATATTAACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAAAGCAGCTGGAATC 5100  
Db 5041 GATTCATATTAACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAAAGCAGCTGGAATC 5100  
Qy 5101 GGAGGATGTTATCTGTCAATGTGTCTGATGAATGAACGGAAGCTTTAGTTAGTATCT 5160  
Db 5101 GGAGGATGTTATCTGTCAATGTGTCTGATGAATGAACGGAAGCTTTAGTTAGTATCT 5160  
Qy 5161 GAGTTGAAGAGTAAGTCTTTTCAATGTAATGCAAAAGTCAAAAGTCAAAAGTCAAAAGTCA 5220  
Db 5161 GAGTTGAAGAGTAAGTCTTTTCAATGTAATGCAAAAGTCAAAAGTCAAAAGTCAAAAGTCA 5220  
Qy 5221 ATTGCGCAAAATGCAAAATGCAAGAAACGCGCTGAGTGTGAGCAACAGTTGCTCATACA 5280  
Db 5221 ATTGCGCAAAATGCAAAATGCAAGAAACGCGCTGAGTGTGAGCAACAGTTGCTCATACA 5280  
Qy 5281 AATTTTGGAAAAATCAATCACTTATAGCTATTGTTAAACAGTAAATTAACAACGCGAAT 5340  
Db 5281 AATTTTGGAAAAATCAATCACTTATAGCTATTGTTAAACAGTAAATTAACAACGCGAAT 5340  
Qy 5341 GATCAAGATGAAGAAAAATATCAATGTAATGCAAAAGTCAAAAGTCAAAAGTCAAAAGTCA 5400  
Db 5341 GATCAAGATGAAGAAAAATATCAATGTAATGCAAAAGTCAAAAGTCAAAAGTCAAAAGTCA 5400  
Qy 5401 GCAAGTCGAGTTGAGAGAGCAAAAGAGCCTCTGTGCAAGAGCTTCTGCAAGTACTACC 5460  
Db 5401 GCAAGTCGAGTTGAGAGAGCAAAAGAGCCTCTGTGCAAGAGCTTCTGCAAGTACTACC 5460  
Qy 5461 TTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTAGAGAA 5520  
Db 5461 TTGAATTAAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTAGAGAA 5520  
Qy 5521 GAAATTAATGGAATTAAGAAAAAGCAATGTTAAATGTTCTAGCTGAAAAATCGAGTCAA 5580  
Db 5521 GAAATTAATGGAATTAAGAAAAAGCAATGTTAAATGTTCTAGCTGAAAAATCGAGTCAA 5580  
Qy 5581 GTGCTCACAAATGCGACAGTGTCTTTCGAGAGCAAGTGACAAAGCTGCAATAGAGACTGGA 5640  
Db 5581 GTGCTCACAAATGCGACAGTGTCTTTCGAGAGCAAGTGACAAAGCTGCAATAGAGACTGGA 5640  
Qy 5641 GTAGCAGTTAATTAATTAACAAAAATCTTGTGACATATTAATAAATAGTCAAAAT 5700  
Db 5641 GTAGCAGTTAATTAATTAACAAAAATCTTGTGACATATTAATAAATAGTCAAAAT 5700  
Qy 5701 GTACGAAATGCTTTGTGTAAGAAAGCAATCTCATTTATTAACAAATGGAATTTGGA 5760  
Db 5701 GTACGAAATGCTTTGTGTAAGAAAGCAATCTCATTTATTAACAAATGGAATTTGGA 5760

Qy 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGTGAACAGTTCTGTAGCAGTGAATTAAGATTGTA 5820  
Db 5761 GCTGAGTTGAGAGCTGAGAGAGCTGAGTGAACAGTTCTGTAGCAGTGAATTAAGATTGTA 5820  
Qy 5821 AATTAATGATGACGAATTAATCAATGCAAAAATCACTGCGAAAGGAAATGTGAGATT 5880  
Db 5821 AATTAATGATGACGAATTAATCAATGCAAAAATCACTGCGAAAGGAAATGTGAGATT 5880  
Qy 5881 ATTACAGAGCTGATGCGGTAAATGCTTAATATACAGAAACAGTGTGAGAGTCCCGT 5940  
Db 5881 ATTACAGAGCTGATGCGGTAAATGCTTAATATACAGAAACAGTGTGAGAGTCCCGT 5940  
Qy 5941 GCAGCAATAGAGGCTCAACAGTGTGATGAATTAACAGATCTACAAAGCATATGTA 6000  
Db 5941 GCAGCAATAGAGGCTCAACAGTGTGATGAATTAACAGATCTACAAAGCATATGTA 6000  
Qy 6001 AAAGATTTACAGATGTTGCTTAAGAAAGAAACAGATGTTATTAATTAATCAAGGCGCA 6060  
Db 6001 AAAGATTTACAGATGTTGCTTAAGAAAGAAACAGATGTTATTAATTAATCAAGGCGCA 6060  
Qy 6061 GTAGTAAAGTGATGAATTAAGTAAATGCAAAATCTTAATTAACGAAGACTTATCAAA 6120  
Db 6061 GTAGTAAAGTGATGAATTAAGTAAATGCAAAATCTTAATTAACGAAGACTTATCAAA 6120  
Qy 6121 AAAAGAAAAATTAATTAATAAAGAAATTTGTACCAATAGTCACTACTCATACTTTA 6180  
Db 6121 AAAAGAAAAATTAATTAATAAAGAAATTTGTACCAATAGTCACTACTCATACTTTA 6180  
Qy 6181 AAATCTTTATTTGCGCAATGCGCTGTGTTACAGAACAGCGGAGTGGCGAAACGTGTAAT 6240  
Db 6181 AAATCTTTATTTGCGCAATGCGCTGTGTTACAGAACAGCGGAGTGGCGAAACGTGTAAT 6240  
Qy 6241 ATCAACAAAGTTTATGAGAAACAGAAAGCTCTGTGAGAAATTTCTATTAATGCAAAA 6300  
Db 6241 ATCAACAAAGTTTATGAGAAACAGAAAGCTCTGTGAGAAATTTCTATTAATGCAAAA 6300  
Qy 6301 CATTAATCTGTAATAATCAGAGATTAACAGAAATCAATCGAGATGTAAGTTCTGTGGT 6360  
Db 6301 CATTAATCTGTAATAATCAGAGATTAACAGAAATCAATCGAGATGTAAGTTCTGTGGT 6360  
Qy 6361 GTTGTGGAATGTAAGAGTGAAGCTTCTTCTGATACCAATATTTATTAAGAAATACC 6420  
Db 6361 GTTGTGGAATGTAAGAGTGAAGCTTCTTCTGATACCAATATTTATTAAGAAATACC 6420  
Qy 6421 AAAGCAAGAGTTGGAAGAAATCAATGTCGATGAGGTTCCGAGAAAGAGCTGAAATT 6480  
Db 6421 AAAGCAAGAGTTGGAAGAAATCAATGTCGATGAGGTTCCGAGAAAGAGCTGAAATT 6480  
Qy 6481 ACAGCAGATTTTAAGCAAGAAATTTCTTTTGAAGTCGAGAGTCGACAGCCGCGGTA 6540  
Db 6481 ACAGCAGATTTTAAGCAAGAAATTTCTTTTGAAGTCGAGAGTCGACAGCCGCGGTA 6540  
Qy 6541 GAGCCGGAAGTGGCAGAACCGTTTCCGTAATCAATTTGCGAGAAAGAGAGTGAAT 6600  
Db 6541 GAGCCGGAAGTGGCAGAACCGTTTCCGTAATCAATTTGCGAGAAAGAGAGTGAAT 6600  
Qy 6601 GTGGAAGAGCAAGAAATTTGTGTAAGAAACCTGAGATTACAGCAAAACGTTATAGTTCT 6660  
Db 6601 GTGGAAGAGCAAGAAATTTGTGTAAGAAACCTGAGATTACAGCAAAACGTTATAGTTCT 6660  
Qy 6661 GTTGCAATTTGAAAAATGCGCAGTCGAGTGTGCTCAAAAGAGCTGGAATTTGAGACACA 6720  
Db 6661 GTTGCAATTTGAAAAATGCGCAGTCGAGTGTGCTCAAAAGAGCTGGAATTTGAGACACA 6720  
Qy 6721 GTGCGATTAACCAAGATGAATCAACACGAGAGCAAGGTGAATAATTTCTAAATTTATG 6780  
Db 6721 GTGCGATTAACCAAGATGAATCAACACGAGAGCAAGGTGAATAATTTCTAAATTTATG 6780  
Qy 6781 ACTGGAACCAAGTTAAGATGAATGAAGAAATGAGATTAATACTAGATGCTGGAATCGGT 6840  
Db 6781 ACTGGAACCAAGTTAAGATGAATGAAGAAATGAGATTAATACTAGATGCTGGAATCGGT 6840

QY 6841 TCAGCCGAGCTGGAATTTCTTGACGCCGAGATCTGAGTGGTTTCTGCAATAATTT 6900  
 Db 6841 TCAGCCGAGCTGGAATTTCTTGACGCCGAGATCTGAGTGGTTTCTGCAATAATTT 6900  
 QY 6901 GCAAAATAGGTGAGAAAGATATGATCATCTTAACCTTTCTACTGATTAAT 6960  
 Db 6901 GCAAAATAGGTGAGAAAGATATGATCATCTTAACCTTTCTACTGATTAAT 6960  
 QY 6961 GTAAAGCTCTTAATTAATTTGAAATTCCTTGACGCGGTGAGAGCCGAGTCTT 7020  
 Db 6961 GTAAAGCTCTTAATTAATTTGAAATTCCTTGACGCGGTGAGAGCCGAGTCTT 7020  
 QY 7021 GCAGAGTTACCGGAGTGGTTCTGTAACTATTAATAGTTCTGTAGTAGCTGAGT 7080  
 Db 7021 GCAGAGTTACCGGAGTGGTTCTGTAACTATTAATAGTTCTGTAGTAGCTGAGT 7080  
 QY 7081 GAGAGATTTACCGGAGTGGTTCTGTAACTATTAATAGTTCTGTAGTAGCTGAGT 7140  
 Db 7081 GAGAGATTTACCGGAGTGGTTCTGTAACTATTAATAGTTCTGTAGTAGCTGAGT 7140  
 QY 7141 AAAAAATTTAAGCAACAGCAGCAATGCAAGATCGAGAGCAGCAATCGAGCCAT 7200  
 Db 7141 AAAAAATTTAAGCAACAGCAGCAATGCAAGATCGAGAGCAGCAATCGAGCCAT 7200  
 QY 7201 GTCTTGGTAATTAATTTTGGAAACAGCTGTGAAGATGAAAAATTTTGAAGAAAGGA 7260  
 Db 7201 GTCTTGGTAATTAATTTTGGAAACAGCTGTGAAGATGAAAAATTTTGAAGAAAGGA 7260  
 QY 7261 ACAGAGTTTAAAAAATTTAGACAGAGTTAAACAAGAACAGATAAAAAGTAAATGAT 7320  
 Db 7261 ACAGAGTTTAAAAAATTTAGACAGAGTTAAACAAGAACAGATAAAAAGTAAATGAT 7320  
 QY 7321 GCTACGAAAAAATCTTACATCAGCAGGTAATTTCTACAGAAATCTTCTGTAAGAGG 7380  
 Db 7321 GCTACGAAAAAATCTTACATCAGCAGGTAATTTCTACAGAAATCTTCTGTAAGAGG 7380  
 QY 7381 GATTAAGAGATCTCAGGAGAGAAATTAAGCCATTTGTGAAGACTCTGATATTAT 7440  
 Db 7381 GATTAAGAGATCTCAGGAGAGAAATTAAGCCATTTGTGAAGACTCTGATATTAT 7440  
 QY 7441 GGAAGAGATCTCAGGAGAGAAATTAAGCCATTTGTGAAGACTCTGATATTAT 7500  
 Db 7441 GGAAGAGATCTCAGGAGAGAAATTAAGCCATTTGTGAAGACTCTGATATTAT 7500  
 QY 7501 GGAAGATCTCAGGAGAGAAATTAAGCCATTTGTGAAGACTCTGATATTAT 7560  
 Db 7501 GGAAGATCTCAGGAGAGAAATTAAGCCATTTGTGAAGACTCTGATATTAT 7560  
 QY 7561 TCCGAGTTACTGTTGAAAAATTTTGTGAAGAGCTGAAAAATGTTAGATCG 7620  
 Db 7561 TCCGAGTTACTGTTGAAAAATTTTGTGAAGAGCTGAAAAATGTTAGATCG 7620  
 QY 7621 GATTAATCAGGAAATGTTGCTTAAACAGATATCAAGTCTGTAGAGCAATTTGGGAA 7680  
 Db 7621 GATTAATCAGGAAATGTTGCTTAAACAGATATCAAGTCTGTAGAGCAATTTGGGAA 7680  
 QY 7681 GAGAGTGCCTATGCAAGATTTAAATCTAAATGGAAGATCAATATCAATTAATAATTC 7740  
 Db 7681 GAGAGTGCCTATGCAAGATTTAAATCTAAATGGAAGATCAATATCAATTAATAATTC 7740  
 QY 7741 AAGCTTTTGAAGAAAAATTTGATGTTATTTGAAGATTAATTCGAAATTTGAGAGGGA 7800  
 Db 7741 AAGCTTTTGAAGAAAAATTTGATGTTATTTGAAGATTAATTCGAAATTTGAGAGGGA 7800  
 QY 7801 GCAAAAGATTAACGTTAGAGGCGTGTAGTCCGAGCCATTAATCAAAAGCAAGAT 7860  
 Db 7801 GCAAAAGATTAACGTTAGAGGCGTGTAGTCCGAGCCATTAATCAAAAGCAAGAT 7860  
 QY 7861 GAAATGAATTCAGAGGTTGAAATTTGAGAGAGATATTTTCAATGAGAAAAATAGAGTACT 7920  
 Db 7861 GAAATGAATTCAGAGGTTGAAATTTGAGAGAGATATTTTCAATGAGAAAAATAGAGTACT 7920  
 QY 7921 AGCCCTTTAAGGAATTTGAGAGAGAAATCAATGTCAAAAGTGAAGAAAAAGAAAGAGTG 7980

Db 7921 AGCCCTTTAAGGAATTTGAGAGAGAAATCAATGTCAAAAGTGAAGAAAAAGAAAGAGTG 7980  
 QY 7981 ACTGCTGAATCTCAAGAGACTTCTGTAGAGCAGTACAGGCGCAGAAATTTCCGA 8040  
 Db 7981 ACTGCTGAATCTCAAGAGACTTCTGTAGAGCAGTACAGGCGCAGAAATTTCCGA 8040  
 QY 8041 GCAAAAGATGCGGAGAGCTCTTAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAA 8100  
 Db 8041 GCAAAAGATGCGGAGAGCTCTTAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAA 8100  
 QY 8101 CATGAGATTAATGTAATATGGAAGCAACATATAAATGAAGTACAGCAATTTCTAAA 8160  
 Db 8101 CATGAGATTAATGTAATATGGAAGCAACATATAAATGAAGTACAGCAATTTCTAAA 8160  
 QY 8161 GCAGTAACAGGTTCTGATTTGGAGAGAGTTGAGTACCAAGCAGAACTACTGCTCA 8220  
 Db 8161 GCAGTAACAGGTTCTGATTTGGAGAGAGTTGAGTACCAAGCAGAACTACTGCTCA 8220  
 QY 8221 GGTAAATCTATGTAAGAGTTGAGAGAGAAATTTGTTCAAGAACAAATCGAATGCA 8280  
 Db 8221 GGTAAATCTATGTAAGAGTTGAGAGAGAAATTTGTTCAAGAACAAATCGAATGCA 8280  
 QY 8281 ATTTCTAAAGTGAAGGTTTGTGATGAATGAATGAATGTAATCTTCTGTAGTATCA 8340  
 Db 8281 ATTTCTAAAGTGAAGGTTTGTGATGAATGAATGTAATGTAATCTTCTGTAGTATCA 8340  
 QY 8341 GGAATGAGAGAGAAATTTCCGAGCAGAGTGAATTAATCTTCAAGCAAAAGTAATCT 8400  
 Db 8341 GGAATGAGAGAGAAATTTCCGAGCAGAGTGAATTAATCTTCAAGCAAAAGTAATCT 8400  
 QY 8401 GAAATCGTAATTCGTTTACGAAGAGATTAATGAATTAATGAATTAATGAATTAAT 8460  
 Db 8401 GAAATCGTAATTCGTTTACGAAGAGATTAATGAATTAATGAATTAATGAATTAAT 8460  
 QY 8461 GAAATCGTAATTCGTTTACGAAGAGATTAATGAATTAATGAATTAATGAATTAAT 8520  
 Db 8461 GAAATCGTAATTCGTTTACGAAGAGATTAATGAATTAATGAATTAATGAATTAAT 8520  
 QY 8521 TTAGCGGTAGCGGTGTCATGCAACAGAGCAAAAGCAATTTACAGATCAAAACAG 8580  
 Db 8521 TTAGCGGTAGCGGTGTCATGCAACAGAGCAAAAGCAATTTACAGATCAAAACAG 8580  
 QY 8581 TTAATCTTCAACCTGTAAATGAGAGAAACGTAATCTCAATCTTGTGCAAAAGCTTTGGCT 8640  
 Db 8581 TTAATCTTCAACCTGTAAATGAGAGAAACGTAATCTCAATCTTGTGCAAAAGCTTTGGCT 8640  
 QY 8641 AAAAATGAAATTTATGAAATGTAAGAACTGAGAGAGCTTGTAGTGCAGAGCAAA 8700  
 Db 8641 AAAAATGAAATTTATGAAATGTAAGAACTGAGAGAGCTTGTAGTGCAGAGCAAA 8700  
 QY 8701 GCAGCGTTGAAATTTATACAAAGATCTACAGAGCAATTTGTTGAGAAATTTGGGA 8760  
 Db 8701 GCAGCGTTGAAATTTATACAAAGATCTACAGAGCAATTTGTTGAGAAATTTGGGA 8760  
 QY 8761 ATTGAAGATTAATTTGAAACGATTCAGAGATTAATGATTTGTAAGATCAACGAGAC 8820  
 Db 8761 ATTGAAGATTAATTTGAAACGATTCAGAGATTAATGATTTGTAAGATCAACGAGAC 8820  
 QY 8821 GGAACCAAGAGAGCTTGTGCAAGAAATGAGTATTTCTGTGAAAAATCAATTTCAAGG 8880  
 Db 8821 GGAACCAAGAGAGCTTGTGCAAGAAATGAGTATTTCTGTGAAAAATCAATTTCAAGG 8880  
 QY 8881 GAAACCAATCATCATTTGAAGATTAAGCAGAAATTTGTAACCGGAAATGTAATGTA 8940  
 Db 8881 GAAACCAATCATCATTTGAAGATTAAGCAGAAATTTGTAACCGGAAATGTAATGTA 8940  
 QY 8941 GATGCTTTGAATGAATCTGATGATGATCTACAGAGAAATGAGTGTGCTATGCTGAAT 9000  
 Db 8941 GATGCTTTGAATGAATCTGATGATGATCTACAGAGAAATGAGTGTGCTATGCTGAAT 9000  
 QY 9001 GGTATTTGAAATGTTGATGATTAATGATGATTAAGAAAAATGTAAGCCAAATTCGA 9060

```
Db 9001 GGTATTGGAATGTTGATGTAATTAATGATTAAGAAAAATGTAGAACCCAAATCCGA 9060
Qy 9061 AGACATGCTATTGTGAAATCTAGTGAAGAAACAAAGATATCAAGCATTTTACAAAGACAAA 9120
Db 9061 AGACATGCTATTGTGAAATCTAGTGAAGAAACAAAGATATCAAGCATTTTACAAAGACAAA 9120
Qy 9121 GTAAATATTCTTGGAAAAAGAGACGCTGACGTGACGTGCAATATCGAATGTACACTT 9180
Db 9121 GTAAATATTCTTGGAAAAAGAGACGCTGACGTGACGTGCAATATCGAATGTACACTT 9180
Qy 9181 TCCATGAGATGATATTTAAATTTTGGCAAGCATATGCTCTTCTCAATTAATACC 9240
Db 9181 TCCATGAGATGATATTTAAATTTTGGCAAGCATATGCTCTTCTCAATTAATACC 9240
Qy 9241 AAAAATTCAAAAAATTAATTTACTTTAGCATCAAGTGAATGCAATGTGAATGTTAT 9300
Db 9241 AAAAATTCAAAAAATTAATTTACTTTAGCATCAAGTGAATGCAATGTGAATGTTAT 9300
Qy 9301 GGGGTGGCTGAAGCAAGAGTGCAGAGCCAAAGCGACAGTTAGTGAAGATCAATA 9360
Db 9301 GGGGTGGCTGAAGCAAGAGTGCAGAGCCAAAGCGACAGTTAGTGAAGATCAATA 9360
Qy 9361 AATAAGACTAATTAATGTTGATTTAGCAGAGAAAAATTAACAGAGGAAACATCATGTA 9420
Db 9361 AATAAGACTAATTAATGTTGATTTAGCAGAGAAAAATTAACAGAGGAAACATCATGTA 9420
Qy 9421 TATGCCGAGTATGATTAATAATTAATTAATTAAGTAAAGCAAAATTCCTAAGGCTATTGCGGAT 9480
Db 9421 TATGCCGAGTATGATTAATAATTAATTAATTAAGTAAAGCAAAATTCCTAAGGCTATTGCGGAT 9480
Qy 9481 GCCAAAGTCATGCTGCAGCTGCTTCGCGCACTGCGCACTATTTGAAAAAATGAAGTAAAA 9540
Db 9481 GCCAAAGTCATGCTGCAGCTGCTTCGCGCACTGCGCACTATTTGAAAAAATGAAGTAAAA 9540
Qy 9541 TTTAATTAATGCAATCCGAGAAATTTTAAATTAATCTGCGCAAGTTGGAAAGGAAAGCTAAT 9600
Db 9541 TTTAATTAATGCAATCCGAGAAATTTTAAATTAATCTGCGCAAGTTGGAAAGGAAAGCTAAT 9600
Qy 9601 AAAAAAAGCGTGGTGAATCTAATCAAGGTAGACTGGTATACGATTAATATATCATGGCAT 9660
Db 9601 AAAAAAAGCGTGGTGAATCTAATCAAGGTAGACTGGTATACGATTAATATATCATGGCAT 9660
Qy 9661 TCTTCTGAAAAAGCATACAAAAAATTGACATATCAATCAAGAGAGAAAAAGGAAA 9720
Db 9661 TCTTCTGAAAAAGCATACAAAAAATTGACATATCAATCAAGAGAGAAAAAGGAAA 9720
Qy 9721 AAATGA 9726
Db 9721 AAATGA 9726
```

## RESULT 2

```
US-10-647-057-8
; Sequence 8, Application US/10647057
; Publication No. US20040047871A1
; GENERAL INFORMATION:
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
; APPLICANT: NAGARAJA, T.
; APPLICANT: STEWART, GEORGE
; APPLICANT: NARAYANAN, SANDEEV
; APPLICANT: CHENGAPPA, M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP
; FILE REFERENCE: PCT-30962
; CURRENT APPLICATION NUMBER: US/10/647,057
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 9726
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-10-647-057-8
```

```
Query Match 100.0%; Score 9726; DB 7; Length 9726;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGCGGCATCAAAAATTAACGTTTCAGAGGCAAGAGAGAGATATCAGATTCTTAAAAA 60
Db 1 ATGAGCGGCATCAAAAATTAACGTTTCAGAGGCAAGAGAGAGATATCAGATTCTTAAAAA 60
Qy 61 GTTTAATGATTTTGGCATTTGTTAATCACTATGACGTTGAGGCTTAATGATCAATC 120
Db 61 GTTTAATGATTTTGGCATTTGTTAATCACTATGACGTTGAGGCTTAATGATCAATC 120
Qy 121 ACCGCACTGAGATTTTGGCAAAAATTAAGAAAAAAGATTAATGTTATGACATTACT 180
Db 121 ACCGCACTGAGATTTTGGCAAAAATTAAGAAAAAAGATTAATGTTATGACATTACT 180
Qy 181 ACAAACAAGATTCAAGGAGAGAACGCTTTTAAACGTTTAAATAGATTGCTTTAACAGAA 240
Db 181 ACAAACAAGATTCAAGGAGAGAACGCTTTTAAACGTTTAAATAGATTGCTTTAACAGAA 240
Qy 241 AATAATATAGCAATCTATATTTTGGGAAAAGAAATAGTACGGGGCTAAATCTTTT 300
Db 241 AATAATATAGCAATCTATATTTTGGGAAAAGAAATAGTACGGGGCTAAATCTTTT 300
Qy 301 AACTTTGTCATGGAAAAATTTGAAGTGAATGGAATTCAACGGAATTCGAGAAATPAA 360
Db 301 AACTTTGTCATGGAAAAATTTGAAGTGAATGGAATTCAACGGAATTCGAGAAATPAA 360
Qy 361 ATGGAAGAAATTTATATTTCTTAAGCTCGAGAGGATGCGAGTAGAAAAATGAGATT 420
Db 361 ATGGAAGAAATTTATATTTCTTAAGCTCGAGAGGATGCGAGTAGAAAAATGAGATT 420
Qy 421 ATCAATGCTGTTCTTTTCACTATATTTCCAAAACAGATGATTTTAAAGAGCTTTG 480
Db 421 ATCAATGCTGTTCTTTTCACTATATTTCCAAAACAGATGATTTTAAAGAGCTTTG 480
Qy 481 GAAGAAGCAAAACATGTGTAAGTTTAAATGAATCAATCCAGTAGATGCAAAAGTAAAA 540
Db 481 GAAGAAGCAAAACATGTGTAAGTTTAAATGAATCAATCCAGTAGATGCAAAAGTAAAA 540
Qy 541 ATTCATTTGAATCCGAATGGAAGCATTAACGTTAAGAGAAAAATCAATGCTGTGAAAGC 600
Db 541 ATTCATTTGAATCCGAATGGAAGCATTAACGTTAAGAGAAAAATCAATGCTGTGAAAGC 600
Qy 601 ATCGTTTATATGCGGCGGATTAATGATGGAATAATCTGCAATACATAAGACGAATT 660
Db 601 ATCGTTTATATGCGGCGGATTAATGATGGAATAATCTGCAATACATAAGACGAATT 660
Qy 661 ACAGATTTTAAAAATTTAGTCAATATATGATGCAATTAATTTCTGCTGACCGAGAT 720
Db 661 ACAGATTTTAAAAATTTAGTCAATATATGATGCAATTAATTTCTGCTGACCGAGAT 720
Qy 721 TTTAAAGCTACAGACAAAATCTGAGATATTAATTTCTTCACTACATGATTTCTCT 780
Db 721 TTTAAAGCTACAGACAAAATCTGAGATATTAATTTCTTCACTACATGATTTCTCT 780
Qy 781 CAAAAAGCTAAGGAAAAAATTTCACTGTTGGAAGAAATGAAGAAATATGTAAGAGA 840
Db 781 CAAAAAGCTAAGGAAAAAATTTCACTGTTGGAAGAAATGAAGAAATATGTAAGAGA 840
Qy 841 AATACCAAGCAATATTTGATCTGATCTGATTTTGAAGCAGATGGAATATATTAATTT 900
Db 841 AATACCAAGCAATATTTGATCTGATCTGATTTTGAAGCAGATGGAATATATTAATTT 900
Qy 901 AGTGGAAAAGCTTACAAATGGGAGATTTATTAAGAAAGAGGAGAAAAAGAACTTATAC 960
Db 901 AGTGGAAAAGCTTACAAATGGGAGATTTATTAAGAAAGAGGAGAAAAAGAACTTATAC 960
Qy 961 ACTGCTTAAATTTATCAGATGTTGAAGCTTCGTTAGAGTAAATTAAGAAAAAGTCTA 1020
Db 961 ACTGCTTAAATTTATCAGATGTTGAAGCTTCGTTAGAGTAAATTAAGAAAAAGTCTA 1020
```

QY 1021 GGAAGAATGTTGACATTACAGCTGAAGCAAGAATTTCTATGATGCAACTTAGTACT 1080  
| | | | |  
Db 1021 GGAAGAATGTTGACATTACAGCTGAAGCAAGAATTTCTATGATGCAACTTAGTACT 1080  
QY 1081 AAGCTTGCAAGCACTCTTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTTAAT 1140  
| | | | |  
Db 1081 AAGCTTGCAAGCACTCTTTAGCTTTGTTACAGGTTCTATTTCTCTATCAATTTTAAT 1140  
QY 1141 GGATTTTAAAGTTTATGACAAAGTACAGTCCGATTCGTTATTTGAGAAAAGTCCAAATC 1200  
| | | | |  
Db 1141 GGATTTTAAAGTTTATGACAAAGTACAGTCCGATTCGTTATTTGAGAAAAGTCCAAATC 1200  
QY 1201 GAAGCAACAGAAAGAAAGCAAAATATTCATTTTACAGTGAAGTAAAGCACTATGGA 1260  
| | | | |  
Db 1201 GAAGCAACAGAAAGAAAGCAAAATATTCATTTTACAGTGAAGTAAAGCACTATGGA 1260  
QY 1261 GCAGCTACTTCTCCATTAAAAATTACCAATTTATTTGAGAAAAGCCAAATGAAAATT 1320  
| | | | |  
Db 1261 GCAGCTACTTCTCCATTAAAAATTACCAATTTATTTGAGAAAAGCCAAATGAAAATT 1320  
QY 1321 CTCAGTATCGAAGCGGATATATTTCTGCAAAAAGTATTCATATGTAATTTGAAGA 1380  
| | | | |  
Db 1321 CTCAGTATCGAAGCGGATATATTTCTGCAAAAAGTATTCATATGTAATTTGAAGA 1380  
QY 1381 GAAGTAAATGGAAGGAGAGCAATATTACTTCAAAATCGAAAATTAATTAATGATGCT 1440  
| | | | |  
Db 1381 GAAGTAAATGGAAGGAGAGCAATATTACTTCAAAATCGAAAATTAATTAATGATGCT 1440  
QY 1441 TCTGTTCTGTTGGAACGATGAGATTCCAATAAGTAGCTCTTCAGTATTTGAGAG 1500  
| | | | |  
Db 1441 TCTGTTCTGTTGGAACGATGAGATTCCAATAAGTAGCTCTTCAGTATTTGAGAG 1500  
QY 1501 GAAAGAGAAAATAAATCTTCCGTCAAGTTGCTAAAGAGCAAAAGTAAATCAGAAACG 1560  
| | | | |  
Db 1501 GAAAGAGAAAATAAATCTTCCGTCAAGTTGCTAAAGAGCAAAAGTAAATCAGAAACG 1560  
QY 1561 GATGATGTAATGAGAGAGGAGCAATTAATTCATTCGAGCTGCTAAAGAGTGA 1620  
| | | | |  
Db 1561 GATGATGTAATGAGAGAGGAGCAATTAATTCATTCGAGCTGCTAAAGAGTGA 1620  
QY 1621 TTGGGGGATGATGATGATGAGGTTGCTGCAAAATATTTCTAATATGCTTCTCC 1680  
| | | | |  
Db 1621 TTGGGGGATGATGATGATGAGGTTGCTGCAAAATATTTCTAATATGCTTCTCC 1680  
QY 1681 CGTATGATGTAATGATGATATCTACATGCCAAGAGCACTAATATGAGGCTCATAC 1740  
| | | | |  
Db 1681 CGTATGATGTAATGATGATATCTACATGCCAAGAGCACTAATATGAGGCTCATAC 1740  
QY 1741 ATTACTAAATAATAGTGTCTGCAACAGGATCTGATTTGGGAATTCCTCAAGTTTATG 1800  
| | | | |  
Db 1741 ATTACTAAATAATAGTGTCTGCAACAGGATCTGATTTGGGAATTCCTCAAGTTTATG 1800  
QY 1801 GATCAGTTTATGATCAGTCACTAATAAATCAATTTTATGATGCAATTAACAGGCTT 1860  
| | | | |  
Db 1801 GATCAGTTTATGATCAGTCACTAATAAATCAATTTTATGATGCAATTAACAGGCTT 1860  
QY 1861 GGAGAGAGCAGTGTCAATGAGGAAATAAGAAATAAGCTAATGATGATGATGATGAT 1920  
| | | | |  
Db 1861 GGAGAGAGCAGTGTCAATGAGGAAATAAGAAATAAGCTAATGATGATGATGATGAT 1920  
QY 1921 GTGTCTGCAACCATGCAATATATATTTCTGCTTCTGCTGCAATATGAGAGAGTGA 1980  
| | | | |  
Db 1921 GTGTCTGCAACCATGCAATATATATTTCTGCTTCTGCTGCAATATGAGAGAGTGA 1980  
QY 1981 AGACTTTCTGAGAGTGAAGGAGTAATGTAAGGCACTTAATGAACTCAAAATCTT 2040  
| | | | |  
Db 1981 AGACTTTCTGAGAGTGAAGGAGTAATGTAAGGCACTTAATGAACTCAAAATCTT 2040  
QY 2041 CGAGGACATTAAGTGAAGTGTGCTGTACGAAAGAGAGAGAGAGAGAGAGAGAGAG 2100  
| | | | |  
Db 2041 CGAGGACATTAAGTGAAGTGTGCTGTACGAAAGAGAGAGAGAGAGAGAGAGAGAG 2100  
QY 2101 GGAATGACAGAGTTTATATGAGAAATTAATAAATATGCTTCTGCTGCAATTCGCGAT 2160  
| | | | |

Db 2101 GGAATGACAGAGTTTATATGAAACTATPAAAAATATGCTTCTGTCGACAAATGCCAT 2160  
| | | | |  
QY 2161 CATGCTGAATTTGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2220  
| | | | |  
Db 2161 CATGCTGAATTTGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2220  
QY 2221 AAAAATCTTCAAAAATGCGCAAGTCTGTTATGATTAATAGAACTTTTAAAGAGCT 2280  
| | | | |  
Db 2221 AAAAATCTTCAAAAATGCGCAAGTCTGTTATGATTAATAGAACTTTTAAAGAGCT 2280  
QY 2281 TTTGAAAAG 2340  
| | | | |  
Db 2281 TTTGAAAAG 2340  
QY 2341 TTATTTGAATGATTTTCAAGAAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
| | | | |  
Db 2341 TTATTTGAATGATTTTCAAGAAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
QY 2401 AGAATGACAAATTTCTTCCGATGGAACCTCAAAAACAGGAACTGCTATAGAAATGCA 2460  
| | | | |  
Db 2401 AGAATGACAAATTTCTTCCGATGGAACCTCAAAAACAGGAACTGCTATAGAAATGCA 2460  
QY 2461 AACTATGTTCAAG 2520  
| | | | |  
Db 2461 AACTATGTTCAAG 2520  
QY 2521 TTTTCAAGAT 2580  
| | | | |  
Db 2521 TTTTCAAGAT 2580  
QY 2581 TATGCAAAATTTTCAACCTTTTACCTTCCGAGCTATGAGAGAGAGAGAGAGAGAGAT 2640  
| | | | |  
Db 2581 TATGCAAAATTTTCAACCTTTTACCTTCCGAGCTATGAGAGAGAGAGAGAGAGAGAT 2640  
QY 2641 GTGGAT 2700  
| | | | |  
Db 2641 GTGGAT 2700  
QY 2701 GAGAGTAATCTTCTGCAAAAAAGATTTAAATATTAAGCTATATTAAGAGAGAGAG 2760  
| | | | |  
Db 2701 GAGAGTAATCTTCTGCAAAAAAGATTTAAATATTAAGCTATATTAAGAGAGAGAG 2760  
QY 2761 GTGAATTTAGTTGAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2820  
| | | | |  
Db 2761 GTGAATTTAGTTGAGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2820  
QY 2821 GGAAGATTAATGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2880  
| | | | |  
Db 2821 GGAAGATTAATGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2880  
QY 2881 GAATTTATCAGAT 2940  
| | | | |  
Db 2881 GAATTTATCAGAT 2940  
QY 2941 TCTTTTAAATGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 3000  
| | | | |  
Db 2941 TCTTTTAAATGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 3000  
QY 3001 ATCAGTAAAGAT 3060  
| | | | |  
Db 3001 ATCAGTAAAGAT 3060  
QY 3061 GCTTTTAAACAGTAAAGATGATCTTCTGTTGAAATGCTGCGGTTCAAGCGGAAATCGGA 3120  
| | | | |  
Db 3061 GCTTTTAAACAGTAAAGATGATCTTCTGTTGAAATGCTGCGGTTCAAGCGGAAATCGGA 3120  
QY 3121 ACGAAAAATGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180  
| | | | |  
Db 3121 ACGAAAAATGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3180  
QY 3181 GCTTCATTTGAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAG 3240  
| | | | |

Db 3181 GCTTCATTGAGATTAATGACGAAAGACAAAGTAATATGATTAAGATTAAGATGATGAA 3240  
Oy 3241 GTAACAGTAATCGCGAATCTTTAGAGTAGATGCAAAAACGACCGGAACATCAACAGT 3300  
Db 3241 GTAACAGTAATCGCGAATCTTTAGAGTAGATGCAAAAACGACCGGAACATCAACAGT 3300  
Oy 3301 ATTTCTGTTGCCGAGGAATTAATAGGTGGAAGTAAACCGAGTGAAGAAAACCGAAA 3360  
Db 3301 ATTTCTGTTGCCGAGGAATTAATAGGTGGAAGTAAACCGAGTGAAGAAAACCGAAA 3360  
Oy 3361 TCAGAAAGAAACCGAGAGGATTTTGTGCAAAATCGGAAACAAAGTGAATCTGTAAAA 3420  
Db 3361 TCAGAAAGAAACCGAGAGGATTTTGTGCAAAATCGGAAACAAAGTGAATCTGTAAAA 3420  
Oy 3421 AATAAATTACGAGATGATGATTCATTAACAGAAAAAATTACAAATTACATTTCTGAA 3480  
Db 3421 AATAAATTACGAGATGATGATTCATTAACAGAAAAAATTACAAATTACATTTCTGAA 3480  
Oy 3481 GGAAGTAAAAAAGCGGGGATCTTCTCGAACGTTTCTCATCTCCGATTAAGACCG 3540  
Db 3481 GGAAGTAAAAAAGCGGGGATCTTCTCGAACGTTTCTCATCTCCGATTAAGACCG 3540  
Oy 3541 TCTTCACTTTGGAGAGCTTCTGGAAGTGTTCCTTCATTAATATTAAGAAACATCT 3600  
Db 3541 TCTTCACTTTGGAGAGCTTCTGGAAGTGTTCCTTCATTAATATTAAGAAACATCT 3600  
Oy 3601 GCTGTCGATGATGAGTAAAGATAATTGAAGGAGCAAAATTAAGAGTGAAGTGAAT 3660  
Db 3601 GCTGTCGATGATGAGTAAAGATAATTGAAGGAGCAAAATTAAGAGTGAAGTGAAT 3660  
Oy 3661 TCTTCTGATCTACTTTTGTGAGACATGGGCGGATCTGCTGCACTTCAGTGAATCAT 3720  
Db 3661 TCTTCTGATCTACTTTTGTGAGACATGGGCGGATCTGCTGCACTTCAGTGAATCAT 3720  
Oy 3721 ATTGGAATGGAATAAGCAATCACTGCTGCTTAACTGAGACGCGCTGCTGAATTAAT 3780  
Db 3721 ATTGGAATGGAATAAGCAATCACTGCTGCTTAACTGAGACGCGCTGCTGAATTAAT 3780  
Oy 3781 ATTCAAAGTAAACCAAGGCTTGGTTAAAAATAGTAGATTCGAAATGCCAATAATTT 3840  
Db 3781 ATTCAAAGTAAACCAAGGCTTGGTTAAAAATAGTAGATTCGAAATGCCAATAATTT 3840  
Oy 3841 AAAGTAAATGCTTTGAGTGAAGAACCTCAAGTAGCAGCAGAGCAGTTTGAAGCAGTT 3900  
Db 3841 AAAGTAAATGCTTTGAGTGAAGAACCTCAAGTAGCAGCAGAGCAGTTTGAAGCAGTT 3900  
Oy 3901 AAAGAAATGGAAGCAAGAAAAATTATCTATTGGGAATCTTCTGCTTCTATCACTTA 3960  
Db 3901 AAAGAAATGGAAGCAAGAAAAATTATCTATTGGGAATCTTCTGCTTCTATCACTTA 3960  
Oy 3961 AAAGAAATGGAAGCAAGAAAAATTATCTATTGGGAATCTTCTGCTTCTATCACTTA 3960  
Db 3961 AAAGAAATGGAAGCAAGAAAAATTATCTATTGGGAATCTTCTGCTTCTATCACTTA 3960  
Oy 3961 GTGAACATGGAAGTTCTGCAAAATCAGAAAAATATACAGTAGCAGAGGAATCTGAAAGC 4020  
Db 3961 GTGAACATGGAAGTTCTGCAAAATCAGAAAAATATACAGTAGCAGAGGAATCTGAAAGC 4020  
Oy 4021 CAAAAAATGATGATGATGTCATCTGCTTATCAAGCGACACCCAAAGTACAGAGCTTTA 4080  
Db 4021 CAAAAAATGATGATGATGTCATCTGCTTATCAAGCGACACCCAAAGTACAGAGCTTTA 4080  
Oy 4081 AATTTACAAAGCTGGAAGATCAAAATGGAACCTGAGGGCTACGTGAATCTGTGCCAAATTA 4140  
Db 4081 AATTTACAAAGCTGGAAGATCAAAATGGAACCTGAGGGCTACGTGAATCTGTGCCAAATTA 4140  
Oy 4141 AACCAACAAGTAAATGCTTCTATTAAGTGGGAGATATACTTAAGTTAATGAGCGGAC 4200  
Db 4141 AACCAACAAGTAAATGCTTCTATTAAGTGGGAGATATACTTAAGTTAATGAGCGGAC 4200  
Oy 4201 GCAAAAAGCTTTTGAACAACAATCAAGTGAATCTGAGTGAAGCAACGGAAGCAATTT 4260  
Db 4201 GCAAAAAGCTTTTGAACAACAATCAAGTGAATCTGAGTGAAGCAACGGAAGCAATTT 4260  
Oy 4261 AGTTCTGAGAGCGGATTAAGAAATTAATCAAGGGGCTGTTCTGTCAATTAAGATGCAAT 4320  
Db 4261 AGTTCTGAGAGCGGATTAAGAAATTAATCAAGGGGCTGTTCTGTCAATTAAGATGCAAT 4320

Oy 4321 GACGTGAAGTACGCTTGAATAAATCTTCCATCGAAGAGCTTAATGAATCAATGTCATT 4380  
Db 4321 GACGTGAAGTACGCTTGAATAAATCTTCCATCGAAGAGCTTAATGAATCAATGTCATT 4380  
Oy 4381 GCCAAAGATGTCAAAGAGTCTGATCTGACAAAAGAAATATCAAGGCTTTACTAAATGA 4440  
Db 4381 GCCAAAGATGTCAAAGAGTCTGATCTGACAAAAGAAATATCAAGGCTTTACTAAATGA 4440  
Oy 4441 AAAGATAAAAAATTTTAAGAAATCGTGTGATTAATGCACTGGAATGGTTATTAACG 4500  
Db 4441 AAAGATAAAAAATTTTAAGAAATCGTGTGATTAATGCACTGGAATGGTTATTAACG 4500  
Oy 4501 AAGGAACAATGAAAAAGCAAAAGAAAAAGAAAGAGACCGGTCATGTAATGCTGCTTTA 4560  
Db 4501 AAGGAACAATGAAAAAGCAAAAGAAAAAGAAAGAGACCGGTCATGTAATGCTGCTTTA 4560  
Oy 4561 TCGGTTCTGGAACGATTAATCCGCTGAGAGAGTACCTATTGCAATCAATCTGTTAAA 4620  
Db 4561 TCGGTTCTGGAACGATTAATCCGCTGAGAGAGTACCTATTGCAATCAATCTGTTAAA 4620  
Oy 4621 AATAAATTAAAGGAAATTAAGTGAAGCAATTAAGAAAGCGGAGAGGATTAATTAATCAT 4680  
Db 4621 AATAAATTAAAGGAAATTAAGTGAAGCAATTAAGAAAGCGGAGAGGATTAATTAATCAT 4680  
Oy 4681 GCGAAACATGTAATGAGAGGCAAAATCATCTACTGTGTTGTGAATGCGGCTTCTGGA 4740  
Db 4681 GCGAAACATGTAATGAGAGGCAAAATCATCTACTGTGTTGTGAATGCGGCTTCTGGA 4740  
Oy 4741 CTGCTATACGAAAGATGCTTTTTCAGAAATGGGAATCTGGAACATGGAAGCTTATCA 4800  
Db 4741 CTGCTATACGAAAGATGCTTTTTCAGAAATGGGAATCTGGAAGCTTATCA 4800  
Oy 4801 AATGACACGATGCAAAAGATGCTTTTTCAGAAATGGGAATCTGGAAGCTTATCA 4860  
Db 4801 AATGACACGATGCAAAAGATGCTTTTTCAGAAATGGGAATCTGGAAGCTTATCA 4860  
Oy 4861 GCAAAATTAATTCATTTCTGGGGTGAATGTTGCGGGAACCAATGCGGTTCTCTTTACG 4920  
Db 4861 GCAAAATTAATTCATTTCTGGGGTGAATGTTGCGGGAACCAATGCGGTTCTCTTTACG 4920  
Oy 4921 GCGGTAGAGAGTCTTTTTCGAAATTAATCTCTTATTAATTAACCTGCTTTATTAACA 4980  
Db 4921 GCGGTAGAGAGTCTTTTTCGAAATTAATCTCTTATTAATTAACCTGCTTTATTAACA 4980  
Oy 4981 GGAACGAAGGTAATCTTTTATGGAAGAAATCAAAAGTCAATGTACAAAGCTTTGAAT 5040  
Db 4981 GGAACGAAGGTAATCTTTTATGGAAGAAATCAAAAGTCAATGTACAAAGCTTTGAAT 5040  
Oy 5041 GATTCGATATTAACAAAGTTTCTGCTGAGAGCGCTCAAGTATTAAGCAGAGCTGAATC 5100  
Db 5041 GATTCGATATTAACAAAGTTTCTGCTGAGAGCGCTCAAGTATTAAGCAGAGCTGAATC 5100  
Oy 5101 GGAAGAAATGTAATCTGTCATCTGCTGTTCTGATGAACCGAAGCTTTATGTAATCTT 5160  
Db 5101 GGAAGAAATGTAATCTGTCATCTGCTGTTCTGATGAACCGAAGCTTTATGTAATCTT 5160  
Oy 5161 GAGTTTGAAGAGTAAGTCTTCAATGTGATGAATCAAAAGATCAAAAGATCAATTAATCA 5220  
Db 5161 GAGTTTGAAGAGTAAGTCTTCAATGTGATGAATCAAAAGATCAAAAGATCAATTAATCA 5220  
Oy 5221 ATTGCGGAAATGCAAAATGAGAAAAAGCGGCTGAGTGAAGCAACGTTGCTCATTAACA 5280  
Db 5221 ATTGCGGAAATGCAAAATGAGAAAAAGCGGCTGAGTGAAGCAACGTTGCTCATTAACA 5280  
Oy 5281 AATAATTGAAAAACAATCAGTTATAGCTAATGTAATAAAAAACGTAATAATTACAACGCGAAT 5340  
Db 5281 AATAATTGAAAAACAATCAGTTATAGCTAATGTAATAAAAAACGTAATAATTACAACGCGAAT 5340  
Oy 5341 GATCAAGTAAAGAAATATCAATGTCAGTCAAAAGATTAATTAATGAACCAATTAATATA 5400  
Db 5341 GATCAAGTAAAGAAATATCAATGTCAGTCAAAAGATTAATTAATGAACCAATTAATATA 5400

QY 5401 GCAGTCGAGATTGAGAGAGCAAAAAGACCTCTGTGCAAGAGGCTTTCGAAGTACTAC 5460  
DB 5401 GCAGTCGAGATTGAGAGAGCAAAAAGACCTCTGTGCAAGAGGCTTTCGAAGTACTAC 5460  
QY 5461 TTGAATAGACAGATTCTTCTCATGTGTGCAAACTGATATTGACAAAGATTGAGGAA 5520  
DB 5461 TTGAATAGACAGATTCTTCTCATGTGTGCAAACTGATATTGACAAAGATTGAGGAA 5520  
QY 5521 GAAATATATGGAATATGGAAGAAAGCAATGTTAATGTTCTAGCTGAAATATCGAGTCAA 5580  
DB 5521 GAAATATATGGAATATGGAAGAAAGCAATGTTAATGTTCTAGCTGAAATATCGAGTCAA 5580  
QY 5581 GTGGTCACAAATGCGACAGTGTCTTCGAGACAAATGAGACAAAGCTGAGTGAAGCTGGA 5640  
DB 5581 GTGGTCACAAATGCGACAGTGTCTTCGAGACAAATGAGACAAAGCTGAGTGAAGCTGGA 5640  
QY 5641 GTAGAGATTATTAATAATTAACAATACTTCTGACATATATAAATATAGTACTCAAAAT 5700  
DB 5641 GTAGAGATTATTAATAATTAACAATACTTCTGACATATATAAATATAGTACTCAAAAT 5700  
QY 5701 GTACGAAATGCTTTGGTAAATAAAGCAAAATCTCATTCATTTAAAAACAATTGGAATTGGA 5760  
DB 5701 GTACGAAATGCTTTGGTAAATAAAGCAAAATCTCATTCATTTAAAAACAATTGGAATTGGA 5760  
QY 5761 GCTGAGATTGAGAGCTGAGAGAGCTGAGAGTGAAGTTCGTGACGTGAATTAAGTTGTA 5820  
DB 5761 GCTGAGATTGAGAGCTGAGAGAGCTGAGAGTGAAGTTCGTGACGTGAATTAAGTTGTA 5820  
QY 5821 AATATATAGATTAAGCAATTAATATCATGCAAAATATCATGCGAAGGAAATGTCGGAAGT 5880  
DB 5821 AATATATAGATTAAGCAATTAATATCATGCAAAATATCATGCGAAGGAAATGTCGGAAGT 5880  
QY 5881 ATTACAGAGTCTGATGCGGTAAATGCTAATTAATGACAGAAACAGTGTCTGAGATGCGCGT 5940  
DB 5881 ATTACAGAGTCTGATGCGGTAAATGCTAATTAATGACAGAAACAGTGTCTGAGATGCGCGT 5940  
QY 5941 GCAGCAATAGAGAGCTCAACAGAGTGAATGAATTAACAGATCTACAAAGACATATGTA 6000  
DB 5941 GCAGCAATAGAGAGCTCAACAGAGTGAATGAATTAACAGATCTACAAAGACATATGTA 6000  
QY 6001 AAAAGTTCTACAGTATGCTTAAGAAGAAACAGATGATTAATTAATTAATTAATTAATTAAT 6060  
DB 6001 AAAAGTTCTACAGTATGCTTAAGAAGAAACAGATGATTAATTAATTAATTAATTAATTAAT 6060  
QY 6061 GTAGATTAAGTGTGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 6120  
DB 6061 GTAGATTAAGTGTGATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 6120  
QY 6121 AAAAGAAAATAAGTAATTAAGTAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 6180  
DB 6121 AAAAGAAAATAAGTAATTAAGTAATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAG 6180  
QY 6181 AAATCTTTATTTGGCAAAATGCGGTGTTACAGACAAAGCGGAGTGGCAGAACTGTATAT 6240  
DB 6181 AAATCTTTATTTGGCAAAATGCGGTGTTACAGACAAAGCGGAGTGGCAGAACTGTATAT 6240  
QY 6241 ATCAACAAGGTTTATGAGAGAAACAGAAAGCTCTTGTAGAAAATTTCTATTTAATGCAAAA 6300  
DB 6241 ATCAACAAGGTTTATGAGAGAAACAGAAAGCTCTTGTAGAAAATTTCTATTTAATGCAAAA 6300  
QY 6301 CATTTATCTGTAAATCAGAGATTAACAGAAATTCAGAGTATAGAGTTCGTGTGAT 6360  
DB 6301 CATTTATCTGTAAATCAGAGATTAACAGAAATTCAGAGTATAGAGTTCGTGTGAT 6360  
QY 6361 GTTGTGTGAAATATGAGAGTATGAGAGTTCGTGTGTATTAATTAATTAATTAATTAATTAAT 6420  
DB 6361 GTTGTGTGAAATATGAGAGTATGAGAGTTCGTGTGTATTAATTAATTAATTAATTAATTAAT 6420  
QY 6421 AAGACAAGAGTTGGAAGAACTCAATGTCTGATGATGAAGTTTGGAGAAAGAGTGAATTT 6480  
DB 6421 AAGACAAGAGTTGGAAGAACTCAATGTCTGATGATGAAGTTTGGAGAAAGAGTGAATTT 6480  
QY 6481 ACAGCAGATTTCAAGCAAGAAATTTCTTTTGAAGTCGAGTTCGACAGACGCGGAGTA 6540

DB 6481 ACAGCAGATTTCAAGCAAGAAATTTCTTTTGAAGTCGAGTTCGACAGACGCGGAGTA 6540  
QY 6541 GGAGCCGAGAGTGGCAGAAACCGTTTCCGTAAATCAATTTTGCAGGAAAGACGGAATGAT 6600  
DB 6541 GGAGCCGAGAGTGGCAGAAACCGTTTCCGTAAATCAATTTTGCAGGAAAGACGGAATGAT 6600  
QY 6601 GTGGAAGAACGAAGATTTTGGTAAATAAAGCTGAGATTAACGCAAAAAGTTATAGTTCT 6660  
DB 6601 GTGGAAGAACGAAGATTTTGGTAAATAAAGCTGAGATTAACGCAAAAAGTTATAGTTCT 6660  
QY 6661 GTTGCAATTGGAATAATGCGCAGTGTGAGTGTGCTCAAAAGAGCTGAAATTTGAGACACA 6720  
DB 6661 GTTGCAATTGGAATAATGCGCAGTGTGAGTGTGCTCAAAAGAGCTGAAATTTGAGACACA 6720  
QY 6721 GTGCGAGTTTACCAAGATGAATCAACACGAGACAGAGTGAATAATTTCTAAATTTATG 6780  
DB 6721 GTGCGAGTTTACCAAGATGAATCAACACGAGACAGAGTGAATAATTTCTAAATTTATG 6780  
QY 6781 ACTGGAACCAAGTTTATGATGATTAATGACAGAAATGAGATTAATTAATTAATTAATTAAT 6840  
DB 6781 ACTGGAACCAAGTTTATGATGATTAATGACAGAAATGAGATTAATTAATTAATTAATTAAT 6840  
QY 6841 TCAAGCCGAGGCTGGAATTTCTGACAGCCGAGATCTGAGTGTGTTCTGTCAATTAATTT 6900  
DB 6841 TCAAGCCGAGGCTGGAATTTCTGACAGCCGAGATCTGAGTGTGTTCTGTCAATTAATTT 6900  
QY 6901 GCAAAATTAAGTAAAGCAAGATATGATATGATATGATATTTTCACTCTTCTACTGATTAAT 6960  
DB 6901 GCAAAATTAAGTAAAGCAAGATATGATATGATATGATATTTTCACTCTTCTACTGATTAAT 6960  
QY 6961 GTAAAAGCTCTTTAATTAATTTCCAAATTTCTTGAACGCGGTGAGAGCCGACAGTCTT 7020  
DB 6961 GTAAAAGCTCTTTAATTAATTTCCAAATTTCTTGAACGCGGTGAGAGCCGACAGTCTT 7020  
QY 7021 GCAGCAGTTTACCGAGTGTGTTCTGTGTAACATTAATTAATGTTCTGTATAGTCTGAGTT 7080  
DB 7021 GCAGCAGTTTACCGAGTGTGTTCTGTGTAACATTAATTAATGTTCTGTATAGTCTGAGTT 7080  
QY 7081 CACATTAATCTGATTTGATTTGATCTTCGTGACAGAAAGTAAATGTAACGGCAAAAGAGAA 7140  
DB 7081 CACATTAATCTGATTTGATTTGATCTTCGTGACAGAAAGTAAATGTAACGGCAAAAGAGAA 7140  
QY 7141 AAAAATTTAAGCAAAACAGACCAAAATGCAAGAAATGCGAGAGCGCAATTCGAGCCAT 7200  
DB 7141 AAAAATTTAAGCAAAACAGACCAAAATGCAAGAAATGCGAGAGCGCAATTCGAGCCAT 7200  
QY 7201 GTCTTGTTGTAATTAATTTTGGAAACAGCTGTAGAAAGTAAAGAAATTTGAAAGAAAGGA 7260  
DB 7201 GTCTTGTTGTAATTAATTTTGGAAACAGCTGTAGAAAGTAAAGAAATTTGAAAGAAAGGA 7260  
QY 7261 ACAGAGTTTAAATACTTTAGACGAGTAAACAAAGAAACAGATTAATAAATGAT 7320  
DB 7261 ACAGAGTTTAAATACTTTAGACGAGTAAACAAAGAAACAGATTAATAAATGAT 7320  
QY 7321 GCTACGAAAATAATCTTCAATCAGCAGTATTTCTACAGAAATCTTCTGTAAAGG 7380  
DB 7321 GCTACGAAAATAATCTTCAATCAGCAGTATTTCTACAGAAATCTTCTGTAAAGG 7380  
QY 7381 GATTAAGAGATATCTCAGGAGAAAGAAATTAAGCAATTTGGAAGCTTCTGATATTAAT 7440  
DB 7381 GATTAAGAGATATCTCAGGAGAAAGAAATTAAGCAATTTGGAAGCTTCTGATATTAAT 7440  
QY 7441 GGAAGAAAATGATGATTAATTAACAGAGGCAAGAAATATATCATCTTCTGCTGTGTTG 7500  
DB 7441 GGAAGAAAATGATGATTAATTAACAGAGGCAAGAAATATATCATCTTCTGCTGTGTTG 7500  
QY 7501 GGAATGCAAGTCTTGTCTTCCGATCAGGAAACAGTGGCAGTTTAAATTTTAAAGAAAT 7560  
DB 7501 GGAATGCAAGTCTTGTCTTCCGATCAGGAAACAGTGGCAGTTTAAATTTTAAAGAAAT 7560  
QY 7561 TCCGAGTTTACTGTTGAAATTTCTTTTGAAGAGCTGAAAGGTAATGTTAGATCG 7620

Db 7561 TCCGAGTTACTGTTGAAAATTCCTTTGTGAAAGCAGCTGAAAAAGTAATGTAGATCG 7620  
Qy 7621 GATATTACAGGAAATGTTGCTTTAACAGCATATCAAGGTCTGTAGAGCATTTGGGAATA 7680  
Db 7621 GATATTACAGGAAATGTTGCTTTAACAGCATATCAAGGTCTGTAGAGCATTTGGGAATA 7680  
Qy 7661 GAGGCTGCTTATGCAGAATTAAATCTTAATGAGATCAAAATATCAGTATTAAAAATCT 7740  
Db 7661 GAGGCTGCTTATGCAGAATTAAATCTTAATGAGATCAAAATATCAGTATTAAAAATCT 7740  
Qy 7741 AAGCTATTAGGAAAAAATTTGATGTTAATGTAATGTAATTCGGAATTTGAGCGGAA 7800  
Db 7741 AAGCTATTAGGAAAAAATTTGATGTTAATGTAATGTAATTCGGAATTTGAGCGGAA 7800  
Qy 7801 GCAAAAAGATTACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAAGCAAGAAT 7860  
Db 7801 GCAAAAAGATTACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAAGCAAGAAT 7860  
Qy 7861 GAAATGAAATTCAGAGGTTGAAAATTGAGAAAGTATTTTCAATGAAGAAAATAGATTAAT 7920  
Db 7861 GAAATGAAATTCAGAGGTTGAAAATTGAGAAAGTATTTTCAATGAAGAAAATAGATTAAT 7920  
Qy 7921 AGCCCTTTTAAAGAAATTTGGAAGAAATCAATGTCAAAGTGGAAAAAAGCAAGATG 7980  
Db 7921 AGCCCTTTTAAAGAAATTTGGAAGAAATCAATGTCAAAGTGGAAAAAAGCAAGATG 7980  
Qy 7981 ACTGCTGATCTCAAGAGCTTCTGTAGAGCAGTAGCGGGCAGAAATTTATTCGGA 8040  
Db 7981 ACTGCTGATCTCAAGAGCTTCTGTAGAGCAGTAGCGGGCAGAAATTTATTCGGA 8040  
Qy 8041 GCAAAAAGATGCCGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100  
Db 8041 GCAAAAAGATGCCGGAAGCTCTTATTTGAAAGTTAGTACAAAATCCGGAAGAAATTTT 8100  
Qy 8101 CATGAGATTAATGCAATATGGAAGCAACACTAAATGAAAGTAAACGCACTTTCTAA 8160  
Db 8101 CATGAGATTAATGCAATATGGAAGCAACACTAAATGAAAGTAAACGCACTTTCTAA 8160  
Qy 8161 GCAAGTAAAGGTTCTGTATTGGGAGAGTGGAGTCACCAAGGCGAAGCTACTGCTCA 8220  
Db 8161 GCAAGTAAAGGTTCTGTATTGGGAGAGTGGAGTCACCAAGGCGAAGCTACTGCTCA 8220  
Qy 8221 GGTAAACTATGCTAGAAAGTTGAGAGAGAAATTTGTCAGAACAATCGATGTAATCA 8280  
Db 8221 GGTAAACTATGCTAGAAAGTTGAGAGAGAAATTTGTCAGAACAATCGATGTAATCA 8280  
Qy 8281 ATTTCTAAAGTGAAGGTTTGAATGAATGAATGAATCTTCTGTAGTATCA 8340  
Db 8281 ATTTCTAAAGTGAAGGTTTGAATGAATGAATGAATCTTCTGTAGTATCA 8340  
Qy 8341 GGAATGAGAGAGAAATTCGCGAGCAGAGTGAATCTTCTACAGCACAAAAGTAACT 8400  
Db 8341 GGAATGAGAGAGAAATTCGCGAGCAGAGTGAATCTTCTACAGCACAAAAGTAACT 8400  
Qy 8401 GAATCCGATGTTGTTTACGAAAGCAAGATTATGAAAATAATGATTACAAAAAATAT 8460  
Db 8401 GAATCCGATGTTGTTTACGAAAGCAAGATTATGAAAATAATGATTACAAAAAATAT 8460  
Qy 8461 ATTTTCAAGAGTCAATGCTCTGCTTTAAATGATCAAAAGATGAAGCAATATGAAATCT 8520  
Db 8461 ATTTTCAAGAGTCAATGCTCTGCTTTAAATGATCAAAAGATGAAGCAATATGAAATCT 8520  
Qy 8521 TTAGCGGTAGCCGAGTGTGATGACACAAGAAACAAGCAATTATGAGATCAAAAG 8580  
Db 8521 TTAGCGGTAGCCGAGTGTGATGACACAAGAAACAAGCAATTATGAGATCAAAAG 8580  
Qy 8581 TTAACCTTACAACTGTAAATGAGAGAAACGTATCTCAACTTCGTGCAAAAAGCTTTGCT 8640  
Db 8581 TTAACCTTACAACTGTAAATGAGAGAAACGTATCTCAACTTCGTGCAAAAAGCTTTGCT 8640  
Qy 8641 AAAAAATGAATAATTATGAAATGTAAAGAACTGAGAGAGCTTAGTCCGAGCGGAAACA 8700  
Db 8641 AAAAAATGAATAATTATGAAATGTAAAGAACTGAGAGAGCTTAGTCCGAGCGGAAACA 8700

Qy 8701 GCAGCCGTTGAAAATTTATACAAAGTACTACAGAGCATTTGTTGACAGAAATTTGGAA 8760  
Db 8701 GCAGCCGTTGAAAATTTATACAAAGTACTACAGAGCATTTGTTGACAGAAATTTGGAA 8760  
Qy 8761 ATTGAGATTAATTTGAAACGATTGCAAGAGATTAATCAGATTGTAAGAGTCAACGAGAC 8820  
Db 8761 ATTGAGATTAATTTGAAACGATTGCAAGAGATTAATCAGATTGTAAGAGTCAACGAGAC 8820  
Qy 8821 GGAACCAAGAGGCTTGTGCGAAAGATGGTATTTCTGGAAGAAAATACAAATTTCAAGG 8880  
Db 8821 GGAACCAAGAGGCTTGTGCGAAAGATGGTATTTCTGGAAGAAAATACAAATTTCAAGG 8880  
Qy 8881 GAAACAAAATATCATCTATGGAAGATTAAGCCAGAAATTTGGAAACCGGAAGTGAATGTA 8940  
Db 8881 GAAACAAAATATCATCTATGGAAGATTAAGCCAGAAATTTGGAAACCGGAAGTGAATGTA 8940  
Qy 8941 GATGCTTTGAATGAATCTGATGATGATCTACAGAGAAAGTGTGCTATGTGGAATT 9000  
Db 8941 GATGCTTTGAATGAATCTGATGATGATCTACAGAGAAAGTGTGCTATGTGGAATT 9000  
Qy 9001 GGTATTGGAATGTTGATGTTAATGTTAATGTTAAGAAAATGTAAGCCAAAATCGGA 9060  
Db 9001 GGTATTGGAATGTTGATGTTAATGTTAATGTTAAGAAAATGTAAGCCAAAATCGGA 9060  
Qy 9061 AGACATGCTATTGTAGAAAACCTACGAAAAACAAGAAATATCAAGCATTTACAAAGCAAAA 9120  
Db 9061 AGACATGCTATTGTAGAAAACCTACGAAAAACAAGAAATATCAAGCATTTACAAAGCAAAA 9120  
Qy 9121 GTAATATTTCTTGGAAAAAGAGACGCTGACGCTGACGCTGACATATGCAATGTAACATT 9180  
Db 9121 GTAATATTTCTTGGAAAAAGAGACGCTGACGCTGACGCTGACATATGCAATGTAACATT 9180  
Qy 9181 TCCATGAGATGATATTTAAATTTTGGCAAGAGATGATGCTCTCAATTAATAACC 9240  
Db 9181 TCCATGAGATGATATTTAAATTTTGGCAAGAGATGATGCTCTCAATTAATAACC 9240  
Qy 9241 AAAAATTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9300  
Db 9241 AAAAATTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9300  
Qy 9301 GGGGTGCTGAAGCAAGAGTGCAGAGCCCAAGCGACAGTTAGTGAAGAAATCAATA 9360  
Db 9301 GGGGTGCTGAAGCAAGAGTGCAGAGCCCAAGCGACAGTTAGTGAAGAAATCAATA 9360  
Qy 9361 AATGAACTAATAATGTTGATTTAGCAGGAAAAATTAACAGAGGGAACATCAATGTA 9420  
Db 9361 AATGAACTAATAATGTTGATTTAGCAGGAAAAATTAACAGAGGGAACATCAATGTA 9420  
Qy 9421 TATGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480  
Db 9421 TATGCCGATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9480  
Qy 9481 GCAAAAAGTCAATGCTGACAGCTGCTCGCAACTCCACTATTGAAAAAAGAAAGTAAAA 9540  
Db 9481 GCAAAAAGTCAATGCTGACAGCTGCTCGCAACTCCACTATTGAAAAAAGAAAGTAAAA 9540  
Qy 9541 TTTAATTAATGGAATCCGAGAAATTTAAATTAATCTGCGCAAGATTGGAAGGAAAGCTAAT 9600  
Db 9541 TTTAATTAATGGAATCCGAGAAATTTAAATTAATCTGCGCAAGATTGGAAGGAAAGCTAAT 9600  
Qy 9601 AAAAAAGTGTGAGATCTAATCAAGTGAATGTAATGATGTAATTAATTAATTAATTAAT 9660  
Db 9601 AAAAAAGTGTGAGATCTAATCAAGTGAATGTAATGATGTAATTAATTAATTAATTAATTAAT 9660  
Qy 9661 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGAGAGAAAAAGGAAA 9720  
Db 9661 TCTTCTGAAAAAGCATCAAAAAATTTGACATATCAATCAAGAGAGAGAAAAAGGAAA 9720  
Qy 9721 AAATGA 9726  
Db 9721 AAATGA 9726

RESULT 3  
US-09-841-786-15 Application US/09841786  
Sequence 15, Application US/09841786  
Patent No. US20020054883A1  
GENERAL INFORMATION:  
APPLICANT: NAGARAJA, T. G.  
APPLICANT: STEWART, GEORGE C.  
APPLICANT: NARAYANAN, SANJEEV K.  
APPLICANT: CHENGAPPA, M. M.  
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
FILE REFERENCE: 30296  
CURRENT APPLICATION NUMBER: US/09/841,786  
CURRENT FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: 09/558,257  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 15  
LENGTH: 11130  
TYPE: DNA  
ORGANISM: Fusobacterium necrophorum  
US-09-841-786-15

Query Match 100.0%; Score 9726; DB 3; Length 11130;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGGGATCAAAATTAAGTTCAGAGGACAGAGAGAGATTCAGATTCTAATAAA 60  
DB 1034 ATGACGGGATCAAAATTAAGTTCAGAGGACAGAGAGAGATTCAGATTCTAATAAA 1093  
QY 61 GTTTTAATGATTTGGATTTGATTAACATTAAGCGGTGAGGCTTAATGATCAATC 120  
DB 1094 GTTTTAATGATTTGGATTTGATTAACATTAAGCGGTGAGGCTTAATGATCAATC 1153  
QY 121 ACCGGAGCTAGAAATTTGGAACAAATAATGAAAAAGATTAATGTTATGACTTACT 180  
DB 1154 ACCGGAGCTAGAAATTTGGAACAAATAATGAAAAAGATTAATGTTATGACTTACT 1213  
QY 181 ACAAACAAGATTCAGAGGAGAGAGCTTTAAGTTAATGATTAATGATTTGCTTAACAGA 240  
DB 1214 ACAAACAAGATTCAGAGGAGAGAGCTTTAAGTTAATGATTAATGATTTGCTTAACAGA 1273  
QY 241 AATAATATGCAATCTATATTTTGGGAAAAAGATAGTACGGGGTAAATATCTTTT 300  
DB 1274 AATAATATGCAATCTATATTTTGGGAAAAAGATAGTACGGGGTAAATATCTTTT 1333  
QY 301 AACTTTGTCATGAAAAATTAAGTAGATGAGGATTAATGATTAATGATTTGCTTTT 360  
DB 1334 AACTTTGTCATGAAAAATTAAGTAGATGAGGATTAATGATTAATGATTTGCTTTT 1393  
QY 361 AATTGAGGAATTTAATTTCTTAAGCTGGAAGGAGTGGCAGTGGAAAAATGAGATT 420  
DB 1394 AATTGAGGAATTTAATTTCTTAAGCTGGAAGGAGTGGCAGTGGAAAAATGAGATT 1453  
QY 421 ATCAATGCTGTTCTTTTCAATCTATATTCACAAAACAAGATGATTTAAGAAGCTTTG 480  
DB 1454 ATCAATGCTGTTCTTTTCAATCTATATTCACAAAACAAGATGATTTAAGAAGCTTTG 1513  
QY 481 GAAGAAGCAAAATGCTAAGTTTAAATGGAATCATTCAGTAGATGAAAAAGTAA 540  
DB 1514 GAAGAAGCAAAATGCTAAGTTTAAATGGAATCATTCAGTAGATGAAAAAGTAA 1573  
QY 541 ATTCCATTAATCCGAATGGAAGCATTAAGTGAAGAAAAATCAATGCTGTGAAGGC 600  
DB 1574 ATTCCATTAATCCGAATGGAAGCATTAAGTGAAGAAAAATCAATGCTGTGAAGGC 1633  
QY 601 ATCGGTTATATGCGCGGATATTAAGTGAAGATATGCAATATCAATGAAGCAGGAAT 660  
DB 1634 ATCGGTTATATGCGCGGATATTAAGTGAAGATATGCAATATCAATGAAGCAGGAAT 1693

QY 661 ACAGATTTAAAAATTTAGTCAATATTAAGTGAATTAATTTGCTGACCGAGAT 720  
DB 1694 ACAGATTTAAAAATTTAGTCAATATTAAGTGAATTAATTTGCTGACCGAGAT 1753  
QY 721 TTAAGAAGTACCAAGCAAAATCTGAGATATTAATTTCTTCACTCAGATGATTTCTCT 780  
DB 1754 TTAAGAAGTACCAAGCAAAATCTGAGATATTAATTTCTTCACTCAGATGATTTCTCT 1813  
QY 781 CAAAAGCTATGGAAGAAAAATTTCACTGTTGGAAGGAATGAAGAAATATGAAGA 840  
DB 1814 CAAAAGCTATGGAAGAAAAATTTCACTGTTGGAAGGAATGAAGAAATATGAAGA 1873  
QY 841 AATAACCAAGCAATATTTGAATCTGATGCTTATTTGGAAGCAGATGCAATATTAATA 900  
DB 1874 AATAACCAAGCAATATTTGAATCTGATGCTTATTTGGAAGCAGATGCAATATTAATA 1933  
QY 901 AGTGGAAAGCTACCAATGGAAGATTTAATAAGAAAGGGAAGAAAACTATTAAC 960  
DB 1934 AGTGGAAAGCTACCAATGGAAGATTTAATAAGAAAGGGAAGAAAACTATTAAC 1993  
QY 961 ACTGCTTTAAGTTATGAGATGTTGGAAGCTCCGTAAGATTAATTAAGAAAAAGTCA 1020  
DB 1994 ACTGCTTTAAGTTATGAGATGTTGGAAGCTCCGTAAGATTAATTAAGAAAAAGTCA 2053  
QY 1021 GGAAGAAGTGTGACATTAACAGTGAAGCAAGAAATTTCTATGATGCAACTTTAGTACT 1080  
DB 2054 GGAAGAAGTGTGACATTAACAGTGAAGCAAGAAATTTCTATGATGCAACTTTAGTACT 2113  
QY 1081 AAGCTTGCAAGCACTCTTTAGCTTTGTTACAGTTCTATTTCTCTATCAATTTAAT 1140  
DB 2114 AAGCTTGCAAGCACTCTTTAGCTTTGTTACAGTTCTATTTCTCTATCAATTTAAT 2173  
QY 1141 GGAATTTTGAATTTATGCAAGTATGTCAGTGTCTTATTTGGAAGAAAGTCCAAATC 1200  
DB 2174 GGAATTTTGAATTTATGCAAGTATGTCAGTGTCTTATTTGGAAGAAAGTCCAAATC 2233  
QY 1201 GAAGCAAGAAAGGAAGCAAAATTTATCTTACAGTGAAGTGAAGCAACTATGGA 1260  
DB 2234 GAAGCAAGAAAGGAAGCAAAATTTATCTTACAGTGAAGTGAAGCAACTATGGA 2293  
QY 1261 GCAGCTACTTCTCAATTAATAATTAACAATTTAATTTGGAAGAAAGCAATGGAACCT 1320  
DB 2294 GCAGCTACTTCTCAATTAATAATTAACAATTTAATTTGGAAGAAAGCAATGGAACCT 2353  
QY 1321 CTGATATCGAGGCGGATATTTCTGCAAAAAGTATTTCCATAGTATCTTGAAGA 1380  
DB 2354 CTGATATCGAGGCGGATATTTCTGCAAAAAGTATTTCCATAGTATCTTGAAGA 2413  
QY 1381 GAAGTAAATGGAAGGAAGGAGAGATTTAATCTTCAAAATCTGAATATCTATGATCT 1440  
DB 2414 GAAGTAAATGGAAGGAAGGAGAGATTTAATCTTCAAAATCTGAATATCTATGATCT 2473  
QY 1441 TCTGTTTCTGTTGGAACGATGAGATTTCAATAAAGTACTTTCAATATTTGATGCT 1500  
DB 2474 TCTGTTTCTGTTGGAACGATGAGATTTCAATAAAGTACTTTCAATATTTGATGCT 2533  
QY 1501 GAAGGAATAATTAATTTCCGTCAAGTTGCTTAAGGACAAAGTGAATCAGAAAAG 1560  
DB 2534 GAAGGAATAATTAATTTCCGTCAAGTTGCTTAAGGACAAAGTGAATCAGAAAAG 2593  
QY 1561 GATGATGTAATGTAAGAGTGAAGGATTTAATCTTCAATGAGCTGCTAATAAGTGA 1620  
DB 2594 GATGATGTAATGTAAGAGTGAAGGATTTAATCTTCAATGAGCTGCTAATAAGTGA 2653  
QY 1621 TTGGGGATAGTGTATAGGGGTTGCTGCTCAATATTTCTAATATATGCTTCTTC 1680  
DB 2654 TTGGGGATAGTGTATAGGGGTTGCTGCTCAATATTTCTAATATATGCTTCTTC 2713  
QY 1681 CGTATATGATGATGATATTTATATGCAAGAAAGCACTTAATATGTAAGGCTCATAC 1740  
DB 2714 CGTATATGATGATGATATTTATATGCAAGAAAGCACTTAATATGTAAGGCTCATAC 2773  
QY 1741 ATTACTAAAAATATGTTCTGCAACAGGATCTGATTTGGGAATTCCAAGTTATGAAT 1800

Db 2774 ATTACTAAATAATAGGTTCTGCAAAAGAGATCGATTTGGGAACTTCCAAAGTTTATGAT 2833  
Qy 1801 GATCAGTTTATGAAATCAGGTATCTAAATCAATTTATAGATGCAATATAACAGCGGTT 1860  
Db 2834 GATCAGTTTATGAAATCAGGTATCTAAATCAATTTATAGATGCAATATAACAGCGGTT 2893  
Qy 1861 GGAGAGACAGTGTCAATGAGGAAATTAAGATAAGCTAACGAACTTATTTAGTCCGT 1920  
Db 2894 GGAGAGACAGTGTCAATGAGGAAATTAAGATAAGCTAACGAACTTATTTAGTCCGT 2953  
Qy 1921 GTGTTCGCAACCATAGCAATCATATAATTTCTGCTTCTGTGGCAATGAGAGAGTGA 1980  
Db 2954 GTGTTCGCAACCATAGCAATCATATAATTTCTGCTTCTGTGGCAATGAGAGAGTGA 3013  
Qy 1981 AGACTTTCTTCAGAGTGGAAAGGAGTAAATGTAAGGCAATTAAATGAAGCTCAAAATCTT 2040  
Db 3014 AGACTTTCTTCAGAGTGGAAAGGAGTAAATGTAAGGCAATTAAATGAAGCTCAAAATCTT 3073  
Qy 2041 CGAGGACTACGTCAAGTGGAAAGTGTGCTGTACGAAAGGAAAGAAAGAACTTATTT 2100  
Db 3074 CGAGGACTACGTCAAGTGGAAAGTGTGCTGTACGAAAGGAAAGAAAGAACTTATTT 3133  
Qy 2101 GGAATTCGACGCTTTTATGAAAATTAATGCTTCTGTGCAATGCGCAT 2160  
Db 3134 GGAATTCGACGCTTTTATGAAAATTAATGCTTCTGTGCAATGCGCAT 3193  
Qy 2161 CATGCTGAATGGTATCCGAAAGGAAATTTGATATCAACAGTGAATTAATTTGAATAT 2220  
Db 3194 CATGCTGAATGGTATCCGAAAGGAAATTTGATATCAACAGTGAATTAATTTGAATAT 3253  
Qy 2221 AAAAAATCCTTCAAAAATGCGAAGTCTGTATTTAGTAATTAAGAACTTTTAAAGAGCT 2280  
Db 3254 AAAAAATCCTTCAAAAATGCGAAGTCTGTATTTAGTAATTTAAAGAGCTTTTAAAGAGCT 3313  
Qy 2281 TTTGGAAAAAGAAACGAAATCTCAGAAATGATCCGAAAGATTTGAATCTATGAAAAA 2340  
Db 3314 TTTGGAAAAAGAAACGAAATCTCAGAAATGATCCGAAAGATTTGAATCTATGAAAAA 3373  
Qy 2341 TTATTAATGATGCTTTTCAAAAAATTTGATGAAAAACGAGCTTTTCTAAATGGTAA 2400  
Db 3374 TTATTAATGATGCTTTTCAAAAAATTTGATGAAAAACGAGCTTTTCTAAATGGTAA 3433  
Qy 2401 AGAATGACAATTAATCTTCCGATGGAACCTTCAAAAAACGAACTGCTATAGAAATTTCA 2460  
Db 3434 AGAATGACAATTAATCTTCCGATGGAACCTTCAAAAAACGAACTGCTATAGAAATTTCA 3493  
Qy 2461 AACTATGTTCAAGGAGAAATGAAAAATTAAGAGAAAAATTAACGAAAGATTTAAAGCT 2520  
Db 3494 AACTATGTTCAAGGAGAAATGAAAAATTAAGAGAAAAATTAACGAAAGATTTAAAGCT 3553  
Qy 2521 TTTTCAGAGAGATGAGTGAATGAATTAAGAACTTTGAATTTTACAGAGTAAAGAAAT 2580  
Db 3554 TTTTCAGAGAGATGAGTGAATGAATTAAGAACTTTGAATTTTACAGAGTAAAGAAAT 3613  
Qy 2581 TATGCAAAATTTTACACTTTTACCTTCCGAGCTAAATGAGAAAGAGATTTCTTCT 2640  
Db 3614 TATGCAAAATTTTACACTTTTACCTTCCGAGCTAAATGAGAAAGAGATTTCTTCT 3673  
Qy 2641 GTGGAGAGAGCTGTTTCTGTGGTGAACAGAGAAATTAAGCAAGTATCCGTTGAAAAA 2700  
Db 3674 GTGGAGAGAGCTGTTTCTGTGGTGAACAGAGAAATTAAGCAAGTATCCGTTGAAAAA 3733  
Qy 2701 GGAGCTAAACTGTCGCAAAAAAAGATTAATTAATTAAGCTAATTAAGCAAAAAACA 2760  
Db 3734 GGAGCTAAACTGTCGCAAAAAAAGATTAATTAATTAAGCTAATTAAGCAAAAAACA 3793  
Qy 2761 GTGAATTTAGTTGAAAATTTGGAATTTGGAACAGTACATCCGAAAGTCAATCCGA 2820  
Db 3794 GTGAATTTAGTTGAAAATTTGGAATTTGGAACAGTACATCCGAAAGTCAATCCGA 3853  
Qy 2821 GGAAGATTTAATGTTCAAAAGATCGAAAAATTCAGCTATCTGTAAGACTTAAGAAAAAGCT 2880

Db 3854 GGAAGATTTAATGTTCAAAAGATCGAAAAATTCAGCTATCTGTAAGACTTAAGAAAAAGCT 3913  
Qy 2881 GAATTAATCAGAGAAAAATTAATTAATGACAGATGCAATTTTATCATGTAGCGGA 2940  
Db 3914 GAATTAATCAGAGAAAAATTAATTAATGACAGATGCAATTTTATCATGTAGCGGA 3973  
Qy 2941 TCTTTTAAATGATGCTCAGGTGGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGT 3000  
Db 3974 TCTTTTAAATGATGCTCAGGTGGAAATGCAATCAATGGAATGGAAGTTATAGTGAAGT 4033  
Qy 3001 ATCAGTAAGGCAAGATTTCCATTTGATGACGAAACATATTTGAAAGCTAATTAAGAAATTT 3060  
Db 4034 ATCAGTAAGGCAAGATTTCCATTTGATGACGAAACATATTTGAAAGCTAATTAAGAAATTT 4093  
Qy 3061 GCTTTTAAACAGTAAGATTAATTAATGATGCTTGTGGAATGTCGCGGTTACGCGGAATCGGA 3120  
Db 4094 GCTTTTAAACAGTAAGATTAATTAATGATGCTTGTGGAATGTCGCGGTTACGCGGAATCGGA 4153  
Qy 3121 ACGAAAAATGCGCGGCTCGGAGTTGCTGTGCGGTTAAATGATTAATTTTCAAAACAA 3180  
Db 4154 ACGAAAAATGCGCGGCTCGGAGTTGCTGTGCGGTTAAATGATTAATTTTCAAAACAA 4213  
Qy 3181 GCTTCATTTGAAGATTAATGACGAAAGCAAGTAAATGATTAAGAAATTAAGATGATGA 3240  
Db 4214 GCTTCATTTGAAGATTAATGACGAAAGCAAGTAAATGATTAAGAAATTAAGATGATGA 4273  
Qy 3241 GTAAACAGTAATGCGGAATCTTTAGAAATGATGCAAAAAACGACCGGAACATCAACAGT 3300  
Db 4274 GTAAACAGTAATGCGGAATCTTTAGAAATGATGCAAAAAACGACCGGAACATCAACAGT 4333  
Qy 3301 ATTTCTGTGCGGAGAAATTAATTAAGTTGAAAGTAAACCGAGTGAAGAAAAACGGA 3360  
Db 4334 ATTTCTGTGCGGAGAAATTAATTAAGTTGAAAGTAAACCGAGTGAAGAAAAACGGA 4393  
Qy 3361 TCAGAAAAAGAACGAGAGGATTTTGTGGCAAAATTCGAAACAAAGTGAATCTGTGTA 3420  
Db 4394 TCAGAAAAAGAACGAGAGGATTTTGTGGCAAAATTCGAAACAAAGTGAATCTGTGTA 4453  
Qy 3421 AATTAATTAATGAGATTAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3480  
Db 4454 AATTAATTAATGAGATTAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4513  
Qy 3481 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3540  
Db 4514 GGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4573  
Qy 3541 TCTTTCAATTTGGAAGCTTCTGGAAGTGTCTTCAATATATTAATTAATTAATTAATTAAT 3600  
Db 4574 TCTTTCAATTTGGAAGCTTCTGGAAGTGTCTTCAATATATTAATTAATTAATTAATTAAT 4633  
Qy 3601 GCTGTCGTAAGATGAGTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
Db 4634 GCTGTCGTAAGATGAGTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4693  
Qy 3661 TCTTCTGATTTCTACTTTTGTGAGCATGGGCGGATCTGCTGCACTTCACTGGAATCAT 3720  
Db 4694 TCTTCTGATTTCTACTTTTGTGAGCATGGGCGGATCTGCTGCACTTCACTGGAATCAT 4753  
Qy 3721 ATTTGGAATGGAATTAATGCAATCATGATGCTGTTAGTGAAGGCGGCTGCTGTAATTAAT 3780  
Db 4754 ATTTGGAATGGAATTAATGCAATCATGATGCTGTTAGTGAAGGCGGCTGCTGTAATTAAT 4813  
Qy 3781 ATTTCAAGTAATTAATCAAGGCTTTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTT 3840  
Db 4814 ATTTCAAGTAATTAATCAAGGCTTTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTT 4873  
Qy 3841 AAAGTAATGCTTTGAGTGAAGAACTCACTAGTACAGCAGAGCAAGTTTGAAGCACTT 3900  
Db 4874 AAAGTAATGCTTTGAGTGAAGAACTCACTAGTACAGCAGAGCAAGTTTGAAGCACTT 4933  
Qy 3901 AAAGAAAGTGAAGCAAGGAAAAAGTTATCTATTTGGAATCTTCTGCTTCTATCACTTA 3960  
Db 4934 AAAGAAAGTGAAGCAAGGAAAAAGTTATCTATTTGGAATCTTCTGCTTCTATCACTTA 4993

OY	3661	GTGAACAATGAAGTTTCTGCAAAATTCAGAAATATATCAAGTAGCAGAGAAATCTGAAGC	4020
Db	4994	GTGAACAATGAAGTTTCTGCAAAATTCAGAAATATATCAAGTAGCAGAGAAATCTGAAGC	5053
OY	4021	CAAAAATGAGTGTGATGTGACATGCTTATCAGCGGACGCCAAGTACAGAGACTTTA	4080
Db	5054	CAAAAATGAGTGTGATGTGACATGCTTATCAGCGGACGCCAAGTACAGAGACTTTA	5113
OY	4081	AATTTCACAGCTGGAAAGTCAATATGGAATCTGAGGGCTACTGACTGTTCGCAATTA	4140
Db	5114	AATTTCACAGCTGGAAAGTCAATATGGAATCTGAGGGCTACTGACTGTTCGCAATTA	5173
OY	4141	AACACAAAGTAAATGCTTCTTATATGCTGTGGAGATATCTAAAGTTATCGAGCGAC	4200
Db	5174	AACACAAAGTAAATGCTTCTTATATGCTGTGGAGATATCTAAAGTTATCGAGCGAC	5233
OY	4201	GCAAAAGCTCTTTTAGCAACCACTCAAGTACCTCTCAGATGACAGCGGAGCAATT	4260
Db	5234	GCAAAAGCTCTTTTAGCAACCACTCAAGTACCTCTCAGATGACAGCGGAGCAATT	5293
OY	4261	AGTTCTGAGCGGGATTGAAATTTATCAAGGGCTGTTTCTGCAATTAAGATGACAT	4320
Db	5294	AGTTCTGAGCGGGATTGAAATTTATCAAGGGCTGTTTCTGCAATTAAGATGACAT	5353
OY	4321	GACGTGAGAGCTAGCGTGTATTAATCTTCCATCGAAGAGCTAAATGAATCATGTCAAT	4380
Db	5354	GACGTGAGAGCTAGCGTGTATTAATCTTCCATCGAAGAGCTAAATGAATCATGTCAAT	5413
OY	4381	GCCAAAGATGTCCAAAGAAAGTTCTGATCTAGCAAAAGAAATTCAGGCTTACTAAATGA	4440
Db	5414	GCCAAAGATGTCCAAAGAAAGTTCTGATCTAGCAAAAGAAATTCAGGCTTACTAAATGA	5473
OY	4441	AAAGATTAATAATATTTAGAAAGATCGTGATTAATACGATGGAATAGTTATTAACG	4500
Db	5474	AAAGATTAATAATATTTAGAAAGATCGTGATTAATACGATGGAATAGTTATTAACG	5533
OY	4501	AAGGAACAATAGAAAAAGCAAAAGAAAAAGAAAGAGCGCATGTGTAATGCTGCTTAA	4560
Db	5534	AAGGAACAATAGAAAAAGCAAAAGAAAAAGAAAGAGCGCATGTGTAATGCTGCTTAA	5593
OY	4561	TCGGTTCGTGAAACGATTAATCCGCTGAGAGATGCTATTGCAATCAATCTGTTAAA	4620
Db	5594	TCGGTTCGTGAAACGATTAATCCGCTGAGAGATGCTATTGCAATCAATCTGTTAAA	5653
OY	4621	AATAAATTTAAAGCGAATTTGATGTGAGAGCAATAAGAAAGCCGAGAGGATAAATTCAT	4680
Db	5654	AATAAATTTAAAGCGAATTTGATGTGAGAGCAATAAGAAAGCCGAGAGGATAAATTCAT	5713
OY	4681	GCGAAACAATGTAATGTTGAGAGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCGGA	4740
Db	5714	GCGAAACAATGTAATGTTGAGAGCAAAATCATCTACTGTTGTTGTAATGCGGCTTCGGA	5773
OY	4741	CTTGCTATCAGCAAAAGATGCTTTTCAGAAATGGGATCTGAGCATGGCAAGACTTATCA	4800
Db	5774	CTTGCTATCAGCAAAAGATGCTTTTCAGAAATGGGATCTGAGCATGGCAAGACTTATCA	5833
OY	4801	AATGACACGATTTGCAAAAGTGTATTAAGAGAAATTTCTGCTGATCTCTTAAATGTGAAC	4860
Db	5834	AATGACACGATTTGCAAAAGTGTATTAAGAGAAATTTCTGCTGATCTCTTAAATGTGAAC	5893
OY	4861	GCAAAATTAATTCATCTTGAGGGGTGAATGTGCGGAGACATTCGCGGTTCTCTTCTACG	4920
Db	5894	GCAAAATTAATTCATCTTGAGGGGTGAATGTGCGGAGACATTCGCGGTTCTCTTCTACG	5953
OY	4921	GCGGTAGAGCTGTTTTCGAAATATATCTTCAATAATAAACCTGCTTTGATTAACA	4980
Db	5954	GCGGTAGAGCTGTTTTCGAAATATATCTTCAATAATAAACCTGCTTTGATTAACA	6013
OY	4981	GGAACGAAGTAAATCTTTTAGTGAAGAAATACAAAAGTCAATGTACAAAGCTTGAAT	5040
Db	6014	GGAACGAAGTAAATCTTTTAGTGAAGAAATACAAAAGTCAATGTACAAAGCTTGAAT	6073

QY	5041	GATTCTCATATTACAAACGTTTCTGCTGAGAGCGCTCCGAAGTAACTAAGCAGCGTGGATC	5100
Db	6074	GATTCCTCATATTACAAACGTTTCTGCTGAGAGCGCTCCGAAGTAACTAAGCAGCGTGGATC	6133
QY	5101	GGAGGAATGCTATCTGCATACGAGGTTCTGATGAATGAACGGAAAGCTTAGTAGATCTC	5160
Db	6134	GGAGGAATGCTATCTGCATACGAGGTTCTGATGAATGAACGGAAAGCTTAGTAGATCTC	6193
QY	5161	GAGTTTGAAGAGTAAAGTCTTTTCAATGTAGATGCAAAAGATCAAAAACATATAATACA	5220
Db	6194	GAGTTTGAAGAGTAAAGTCTTTTCAATGTAGATGCAAAAGATCAAAAACATATAATACA	6253
QY	5221	ATTGCGGAAATGCAAAATGAGGAAAAGCGCTGCGAGTTGAGACACAAGTTGCTCATACA	5280
Db	6254	ATTGCGGAAATGCAAAATGAGGAAAAGCGCTGCGAGTTGAGACACAAGTTGCTCATACA	6313
QY	5281	AAATATGGAAACAATCAGTTATAGCTATTGTAAAAAACGTAAATTTACACAGCGCAT	5340
Db	6314	AAATATGGAAACAATCAGTTATAGCTATTGTAAAAAACGTAAATTTACACAGCGCAT	6373
QY	5341	GATCAAGATAGAAAAAATATCATATGTGACGTGCAAAAAGATTATACATGACCAATACTATA	5400
Db	6374	GATCAAGATAGAAAAAATATCATATGTGACGTGCAAAAAGATTATACATGACCAATACTATA	6433
QY	5401	GCAGTGCGAGTTGGAGAGGCAAAAGAGCCTCTGTGCAGAGAGCTTCTGCAGATCTAC	5460
Db	6434	GCAGTGCGAGTTGGAGAGGCAAAAGAGCCTCTGTGCAGAGAGCTTCTGCAGATCTAC	6493
QY	5461	TTGGAATAGACAGTTTCTTCTCATGTGATCTCAACGTATNTGCAAAAGATTAGAGGAA	5520
Db	6494	TTGGAATAGACAGTTTCTTCTCATGTGATCTCAACGTATNTGCAAAAGATTAGAGGAA	6553
QY	5521	GAAAAATATGAGAAATTAAGAAAAAGGCAAAATGTTATGTTCTAGCTGCAAAATACGAGTCAA	5580
Db	6554	GAAAAATATGAGAAATTAAGAAAAAGGCAAAATGTTATGTTCTAGCTGCAAAATACGAGTCAA	6613
QY	5581	GTGCTGACAAATGCGACAGTGCCTTTCGAGCAAGTGAACAAGTGCAGTGGAGCTGGA	5640
Db	6614	GTGCTGACAAATGCGACAGTGCCTTTCGAGCAAGTGAACAAGTGCAGTGGAGCTGGA	6673
QY	5641	GTAGCAGTTAATTAATTAACAACAATACTTCTGCACTATTAATAAATATGATCTCAAAAT	5700
Db	6674	GTAGCAGTTAATTAATTAACAACAATACTTCTGCACTATTAATAAATATGATCTCAAAAT	6733
QY	5701	GTACGAAATGCTTGGTGTAAAGCAAAATCTCATTCATCTATTAAACAATTTGGAATTGGA	5760
Db	6734	GTACGAAATGCTTGGTGTAAAGCAAAATCTCATTCATCTATTAAACAATTTGGAATTGGA	6793
QY	5761	GCTGAGTTGAGCTGAGAGAGCTGAGATGACAAGTCTGTAGCAGTGAATTAAGATTGTA	5820
Db	6794	GCTGAGTTGAGCTGAGAGAGCTGAGATGACAAGTCTGTAGCAGTGAATTAAGATTGTA	6853
QY	5821	AATTAATACGATAGCAGAAATTAAATCATGCAAAAATCACTGCGAAAGGAAATGTCCGAGTT	5880
Db	6854	AATTAATACGATAGCAGAAATTAAATCATGCAAAAATCACTGCGAAAGGAAATGTCCGAGTT	6913
QY	5881	ATTACAGAGTCTGAATGCGGTAAATGCTTAAATTATGAGGAACAGTGTCTGAGTGGCCCGT	5940
Db	6914	ATTACAGAGTCTGAATGCGGTAAATGCTTAAATTATGAGGAACAGTGTCTGAGTGGCCCGT	6973
QY	5941	GCAGCAATAGAGCCTCAACCAAGTGTGATGAATTAACAGATCTCAAAAAGCATATGTA	6000
Db	6974	GCAGCAATAGAGCCTCAACCAAGTGTGATGAATTAACAGATCTCAAAAAGCATATGTA	7033
QY	6001	AAAGATTCTACAGTATTTGCTTAAAGAGAAACAGATGTTATTTATTACTACCAAGGGGAA	6066
Db	7034	AAAGATTCTACAGTATTTGCTTAAAGAGAAACAGATGTTATTTATTACTACCAAGGGGAA	7093
QY	6061	GTAGATTAAGTGTGATTAAGTAACTTCAAAATCTTAATTTAAAGAAAGCTTATACAA	6120
Db	7094	GTAGATTAAGTGTGATTAAGTAACTTCAAAATCTTAATTTAAAGAAAGCTTATACAA	7153
QY	6121	AAAAAATAAGTAAATTAATAAAAGATTTGTTACCAATAGTCACTACTCATATCTTTA	6180

Db 7154 AAAAAAAAAAATAGATATAAAAAGATTGTTACCAATAGTTCAGCTACATCTTTA 7213  
Qy 6181 AATCTTTATTTGGCAATATCCGCTGGTTTCAGGACAAAGCGAGTGGCAGAACTGTAT 6240  
Db 7214 AATCTTTATTTGGCAATATCCGCTGGTTTCAGGACAAAGCGAGTGGCAGAACTGTAT 7273  
Qy 6241 ATCAACAAGTTTATGAGAAAACAGAAGCTCTGTAGAAAAATTCATATTAATGCAAAA 6300  
Db 7274 ATCAACAAGTTTATGAGAAAACAGAAGCTCTGTAGAAAAATTCATATTAATGCAAAA 7333  
Qy 6301 CATTTATTTCTGTAATTCAGAGATTTACAGAAATTCATTCGAGTGTAGGTTCTGTGCT 6360  
Db 7334 CATTTATTTCTGTAATTCAGAGATTTACAGAAATTCATTCGAGTGTAGGTTCTGTGCT 7393  
Qy 6361 GTTGTGGAATATGATGAGATAGAGCTTCTGTATCAATATTAATTAAGAAATACC 6420  
Db 7394 GTTGTGGAATATGATGAGATAGAGCTTCTGTATCAATATTAATTAAGAAATACC 7453  
Qy 6421 AAGACAAGATTGAAAAAACTACAAATGTCATGATGAAGTTTCGAGAAAGCTGAAT 6480  
Db 7454 AAGACAAGATTGAAAAAACTACAAATGTCATGATGAAGTTTCGAGAAAGCTGAAT 7513  
Qy 6481 ACAGCAGATTCTAACAGAGAAATTTCTCTTTTGAATCGAGTCGACAGCCGGGTA 6540  
Db 7514 ACAGCAGATTCTAACAGAGAAATTTCTCTTTTGAATCGAGTCGAGTCGACAGCCGGGTA 7573  
Qy 6541 GGAGCCGAGTGGCAGAACCGTTCCGTAATCAATTTGCAAGAAAGCGGAATAGAT 6600  
Db 7574 GGAGCCGAGTGGCAGAACCGTTCCGTAATCAATTTGCAAGAAAGCGGAATAGAT 7633  
Qy 6601 GTGAGAAGAACAGAAATTTGTGTAAAAAAAGCTGAGATTACAGCAAAACGTTATAGTCT 6660  
Db 7634 GTGAGAAGAACAGAAATTTGTGTAAAAAAAGCTGAGATTACAGCAAAACGTTATAGTCT 7693  
Qy 6661 GTTGCAATTTGAAAAATGCCGAGTCGAGTGGCTGCAAAAGAGCTGGAATTTGACAGCA 6720  
Db 7694 GTTGCAATTTGAAAAATGCCGAGTCGAGTGGCTGCAAAAGAGCTGGAATTTGACAGCA 7753  
Qy 6721 GTGAGAGTTTACCAAGATGATCAAAACAGAGCAAGAGTGAATAATTTCTAAATTTATG 6780  
Db 7754 GTGAGAGTTTACCAAGATGATCAAAACAGAGCAAGAGTGAATAATTTCTAAATTTATG 7813  
Qy 6781 ACTCGAAAACAAGTTAGATGTATATAGCAGAAAAATGAGATTAATTCAGATTCGATTCGGT 6840  
Db 7814 ACTCGAAAACAAGTTAGATGTATATAGCAGAAAAATGAGATTAATTCAGATTCGATTCGGT 7873  
Qy 6841 TCAGCCGAGCTGGAATTTCTTCAGCCGAGATATCTGAGTGGTTTCTGTCAATTAATTT 6900  
Db 7874 TCAGCCGAGCTGGAATTTCTTCAGCCGAGATATCTGAGTGGTTTCTGTCAATTAATTT 7933  
Qy 6901 GCAAAATTAAGGTAGAAACAGATATGATCATAGTACTTTCACTCTTCAATGATTAAT 6960  
Db 7934 GCAAAATTAAGGTAGAAACAGATATGATCATAGTACTTTCACTCTTCAATGATTAAT 7993  
Qy 6961 GTAAAAAGCTCTTAATTAATTTTGAATTTCTTGAAGCCGATGAGAGCCGACAGTCTT 7020  
Db 7994 GTAAAAAGCTCTTAATTAATTTTGAATTTCTTGAAGCCGATGAGAGCCGACAGTCTT 8053  
Qy 7021 GCAGCAGTTTACCGAGTGGTTTCTGTAAACATTAATATAGTTCTGTATAGCTCGAGTT 7080  
Db 8054 GCAGCAGTTTACCGAGTGGTTTCTGTAAACATTAATATAGTTCTGTATAGCTCGAGTT 8113  
Qy 7081 CACAATTAATCTGATTTGATTCCTCGTAAGAAAAAATGTAATCGGCAAAAGAGAA 7140  
Db 8114 CACAATTAATCTGATTTGATTCCTCGTAAGAAAAAATGTAATCGGCAAAAGAGAA 8173  
Qy 7141 AAAAAATTTAAGCAAAACAGCAGCAAAATGAGAGATTCGAGAGCAGCAATCGAGCCAT 7200  
Db 8174 AAAAAATTTAAGCAAAACAGCAGCAAAATGAGAGATTCGAGAGCAGCAATCGAGCCAT 8233  
Qy 7201 GTCTTGTAATTAATTTTGGAAACAGCTGTAGAAAGATGAAAAATTTCTGAAGAAAAAGA 7260

Db 8234 GTCTTGTAATTAATTTTGGAAACAGCTGTAGAAAGATGAAAAAATTTCTGAAGAAAAAGA 8293  
Qy 7261 ACAGAAAGTTTAAAAAATTTTGAAGCAAGTTTAAACAAGAACAGATTAATAAATGAT 7320  
Db 8294 ACAGAAAGTTTAAAAAATTTTGAAGCAAGTTTAAACAAGAACAGATTAATAAATGAT 8353  
Qy 7321 GCTACGAAAAAATCTTAACAATCAGCAGATATTTTCTACAGAAATCTTCTGTAAAGCG 7380  
Db 8354 GCTACGAAAAAATCTTAACAATCAGCAGATATTTTCTACAGAAATCTTCTGTAAAGCG 8413  
Qy 7381 GATAGAGAGTACTCAGAGGAGAAAGAAATTAACCAATTTGAGAGACTTCTGATTAATTT 7440  
Db 8414 GATAGAGAGTACTCAGAGGAGAAAGAAATTAACCAATTTGAGAGACTTCTGATTAATTT 8473  
Qy 7441 GGAATAAATGTAGATATTTATCAACAGAGAACAGAAATTAATCACTTCTGATGATTTG 7500  
Db 8474 GGAATAAATGTAGATATTTATCAACAGAGAACAGAAATTAATCACTTCTGATGATTTG 8533  
Qy 7501 GGAATCGAGGCTTTCCTCCGATCAGAAACAGTGGCAGTTACAAATATTTAAAAAAT 7560  
Db 8534 GGAATCGAGGCTTTCCTCCGATCAGAAACAGTGGCAGTTACAAATATTTAAAAAAT 8593  
Qy 7561 TCCGAGATTACTGTTGAAAAATTTCTTTTGTGAAGACGCTGAAAAAGTTAATGATCG 7620  
Db 8594 TCCGAGATTACTGTTGAAAAATTTCTTTTGTGAAGACGCTGAAAAAGTTAATGATCG 8653  
Qy 7621 GATATTTACAGAAATGTTCCTTTAAACAGCATATCAAGGCTCTGTAGAGCATTTGGAAAT 7680  
Db 8654 GATATTTACAGAAATGTTCCTTTAAACAGCATATCAAGGCTCTGTAGAGCATTTGGAAAT 8713  
Qy 7681 GGAGCTGCTATGCAAAATTTAATTTCTAATGAAAGATCAAAATATCAGTATTAATAATCT 7740  
Db 8714 GGAGCTGCTATGCAAAATTTAATTTCTAATGAAAGATCAAAATATCAGTATTAATAATCT 8773  
Qy 7741 AAGCTATTTAGAAAAAATATTTGATGTTATTTGTAAGTAATGGAATTTGAGACCGAA 7800  
Db 8774 AAGCTATTTAGAAAAAATATTTGATGTTATTTGTAAGTAATGGAATTTGAGACCGAA 8833  
Qy 7801 GCAAAAAGATTAACCGTAGAGAGCGGTAGCTCCGAGCAGTATCTCAAAAAGCAAGAT 7860  
Db 8834 GCAAAAAGATTAACCGTAGAGAGCGGTAGCTCCGAGCAGTATCTCAAAAAGCAAGAT 8893  
Qy 7861 GAAATGAATTCAGAGTTGAAATTTGAGAGATATTTTCAATGAAGAAATAGAGTAACT 7920  
Db 8894 GAAATGAATTCAGAGTTGAAATTTGAGAGATATTTTCAATGAAGAAATAGAGTAACT 8953  
Qy 7921 AGCCCTTCTAAAGAAATTTGAGAGAAATCAATGTCAGAGTGAAGAAAGAAACAGATG 7980  
Db 8954 AGCCCTTCTAAAGAAATTTGAGAGAAATCAATGTCAGAGTGAAGAAAGAAACAGATG 9013  
Qy 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGAGAGGCGAGAAATTAATTCGAA 8040  
Db 9014 ACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGAGAGGCGAGAAATTAATTCGAA 9073  
Qy 8041 GCAAAAAGTCCGAGAGCTCTTAATTTGAAAGTTAGTACAAATCCGAGAAAGATATTTT 8100  
Db 9074 GCAAAAAGTCCGAGAGCTCTTAATTTGAAAGTTAGTACAAATCCGAGAAAGATATTTT 9133  
Qy 8101 CATGCAATTAATGTGAATATGGAAGCAACATTAATTAAGAAATGAAACAGCTTTCTAA 8160  
Db 9134 CATGCAATTAATGTGAATATGGAAGCAACATTAATTAAGAAATGAAACAGCTTTCTAA 9193  
Qy 8161 GCAGTAACAGGTTCTGTATTTGAGAGAGTTGAGTGCACAAAGGCAAGAACTACTGCTCA 8220  
Db 9194 GCAGTAACAGGTTCTGTATTTGAGAGAGTTGAGTGCACAAAGGCAAGAACTACTGCTCA 9253  
Qy 8221 GGTAAAACTATGATGAGAGTTGAGAGAAATTTGTTCAAGAACAAATCGATTGAATGCA 8280  
Db 9254 GGTAAAACTATGATGAGAGTTGAGAGAAATTTGTTCAAGAACAAATCGATTGAATGCA 9313  
Qy 8281 ATTTCTAAAGTGAAGAGTTTGGATGAGATTAAGTAACTGCTAAATCTTCTGTAGTATCA 8340  
Db 9314 ATTTCTAAAGTGAAGAGTTTGGATGAGATTAAGTAACTGCTAAATCTTCTGTAGTATCA 9373

```
QY 8341 GGAATGAGAGAGATTCCCGAGCAGAGTGAATCTTCTACAGCACAAAAGTAACTCT 8400
| | | | |
Db 9374 GGAATGAGAGAGATTCCCGAGCAGAGTGAATCTTCTACAGCACAAAAGTAACTCT 9433
QY 8401 GAAATCCGATTCGTTTACGAAAGCAAGATTATGAAATTAATGATTACACAAAAAATAT 8460
| | | | |
Db 9434 GAAATCCGATTCGTTTACGAAAGCAAGATTATGAAATTAATGATTACACAAAAAATAT 9493
QY 8461 ATTTGAGAGTCAATGCTCTTCTTAATGATACAAAGAAATGAAGCAATATAGAACTCT 8520
| | | | |
Db 9494 ATTTGAGAGTCAATGCTCTTCTTAATGATACAAAGAAATGAAGCAATATAGAACTCT 9553
QY 8521 TTAGCGTAGCCGCGTGTSCATGACACAGAAACAAACAAAGCAATTTACAGATCAAAACAG 8580
| | | | |
Db 9554 TTAGCGTAGCCGCGTGTSCATGACACAGAAACAAACAAAGCAATTTACAGATCAAAACAG 9613
QY 8581 TTAACTTCTACAACTGTTAAATGAGAGAAACGATCTCAACTTCGTGCAAAAGCTTTGGCT 8640
| | | | |
Db 9614 TTAACTTCTACAACTGTTAAATGAGAGAAACGATCTCAACTTCGTGCAAAAGCTTTGGCT 9673
QY 8641 AAAAATGAAATTTATGAAATGTAAGGAACTGAGAGAGCCTTAGTCGAGCGGAAACA 8700
| | | | |
Db 9674 AAAAATGAAATTTATGAAATGTAAGGAACTGAGAGAGCCTTAGTCGAGCGGAAACA 9733
QY 8701 GCAGCGTTGAAATTTATACAAAGATACACAGAGCAATTCGTTGACAGAAATTTGGAA 8760
| | | | |
Db 9734 GCAGCGTTGAAATTTATACAAAGATACACAGAGCAATTCGTTGACAGAAATTTGGAA 9793
QY 8761 ATTGAGATTAATTTGAAACCGATTCCAGAGATTAATGATGTAAGTCAACGAGAC 8820
| | | | |
Db 9794 ATTGAGATTAATTTGAAACCGATTCCAGAGATTAATGATGTAAGTCAACGAGAC 9853
QY 8821 GGAACCAAGAGAGTCTTGTCCGAAAGAAATGATTTCTGTGAAATTAACAATTTCAAGG 8880
| | | | |
Db 9854 GGAACCAAGAGAGTCTTGTCCGAAAGAAATGATTTCTGTGAAATTAACAATTTCAAGG 9913
QY 8881 GAAACAAATCATCATTCATGAGATTAAGCCAGAAATTTGTAACCGGAAAGTGAATGTA 8940
| | | | |
Db 9914 GAAACAAATCATCATTCATGAGATTAAGCCAGAAATTTGTAACCGGAAAGTGAATGTA 9973
QY 8941 GATGCTTTGATGAACCTTGATGATGATCTACAGAGAAAAAGTGTGCTAATGTTGAAT 9000
| | | | |
Db 9974 GATGCTTTGATGAACCTTGATGATGATCTACAGAGAAAAAGTGTGCTAATGTTGAAT 10033
QY 9001 GGTATTGGAATGTTGATGATTAATATGATTAAGAAAAATGTAAGCCAAAAATTCGA 9060
| | | | |
Db 10034 GGTATTGGAATGTTGATGATTAATATGATTAAGAAAAATGTAAGCCAAAAATTCGA 10093
QY 9061 AGACATGCTATTGTAGAAACTACGAGAAACAGAAATATCAAGCATTTTCAAGAGCAAA 9120
| | | | |
Db 10094 AGACATGCTATTGTAGAAACTACGAGAAACAGAAATATCAAGCATTTTCAAGAGCAAA 10153
QY 9121 GTPAATATTCTTGGAAAAAGAGACGCTGACGCTGCAATATCGAATGTACACATT 9180
| | | | |
Db 10154 GTPAATATTCTTGGAAAAAGAGACGCTGACGCTGCAATATCGAATGTACACATT 10213
QY 9181 TCCATATGATGATTTAAAAATTTGGCAAGCAGATATGCAATCTTCAATTAATTAAC 9240
| | | | |
Db 10214 TCCATATGATGATTTAAAAATTTGGCAAGCAGATATGCAATCTTCAATTAATTAAC 10273
QY 9241 AAAAATTCAAAAAATTAATTAATCTTTAGCATCAAGTATGAAATGTAATGTTCAAT 9300
| | | | |
Db 10274 AAAAATTCAAAAAATTAATTAATCTTTAGCATCAAGTATGAAATGTAATGTTCAAT 10333
QY 9301 GGGGTGCTGAGCAAGAGGTGAGAGCCAAAGCAGATTATGTTAAAGATCAAAAT 9360
| | | | |
Db 10334 GGGGTGCTGAGCAAGAGGTGAGAGCCAAAGCAGATTATGTTAAAGATCAAAAT 10393
QY 9361 AATAGAACTTAATATGTTGATTTAGCAGAGAAAAATTTAAACAGAGGAAAAATCAATGTA 9420
| | | | |
Db 10394 AATAGAACTTAATATGTTGATTTAGCAGAGAAAAATTTAAACAGAGGAAAAATCAATGTA 10453
```

```
QY 9421 TATCCGGATATGATTAATAATTAATATAGTAAGACAAATTTCAAGCTATTCCGAT 9480
| | | | |
Db 10454 TATCCGGATATGATTAATAATTAATATAGTAAGACAAATTTCAAGCTATTCCGAT 10513
QY 9481 GCCAAAAGTCATGCTGACGCTCTTGGCAATGCCACTATTTGAAAAAATGAAGTAA 9540
| | | | |
Db 10514 GCCAAAAGTCATGCTGACGCTCTTGGCAATGCCACTATTTGAAAAAATGAAGTAA 10573
QY 9541 TTTTAATATGATCCGAGAAATTTAAAAATTAATTCGCAAGATTGGAAGGGAAGCTAAT 9600
| | | | |
Db 10574 TTTTAATATGATCCGAGAAATTTAAAAATTAATTCGCAAGATTGGAAGGGAAGCTAAT 10633
QY 9601 AAAAAACGTGATGATCTAATCAGTAGACTGTATACGATTAATATACATGCGAT 9660
| | | | |
Db 10634 AAAAAACGTGATGATCTAATCAGTAGACTGTATACGATTAATATACATGCGAT 10693
QY 9661 TCTTCTGAAAAAGATCAAAAAATTTGACATATCAATCAAGAGAGGAAAAAGGAAA 9720
| | | | |
Db 10694 TCTTCTGAAAAAGATCAAAAAATTTGACATATCAATCAAGAGAGGAAAAAGGAAA 10753
QY 9721 AATATGA 9726
| | | | |
Db 10754 AATATGA 10759

RESULT 4
US-10-647-057-15
; Sequence 15, Application US/10647057
; Publication No. US20040047871A1
; GENERAL INFORMATION:
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
; APPLICANT: MACRABA, T.
; APPLICANT: STEWART, GEORGE
; APPLICANT: NARAYANAN, SANJEEV
; APPLICANT: CHENGAPPA, M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: PCT-30962
; CURRENT APPLICATION NUMBER: US/10/647.057
; CURRENT FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 11130
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-10-647-057-15

Query Match 100.0%; Score 9726; DB 7; Length 11130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCGGATCAAAATTAACGTTGAGAGACAAAGAGAGATATCAGATTCTAAAAA 60
| | | | |
Db 1034 ATGAGCGGATCAAAATTAACGTTGAGAGACAAAGAGAGATATCAGATTCTAAAAA 1093
QY 61 GTTTTAATGATTTGGGATGTTGATTAACACTATGACGCTGAGGCTTAATGATCAATC 120
| | | | |
Db 1094 GTTTTAATGATTTGGGATGTTGATTAACACTATGACGCTGAGGCTTAATGATCAATC 1153
QY 121 ACCGGACATGAAATTTTGGAAACAAAAATGAAAAAAGATTAATGATTAATGACTTACT 180
| | | | |
Db 1154 ACCGGACATGAAATTTTGGAAACAAAAATGAAAAAAGATTAATGATTAATGACTTACT 1213
QY 181 ACAACAAGATTCAGAGGAGAACGCTTTTAAACAGTTTAAATAGATTGCTTTAACAGAA 240
| | | | |
Db 1214 ACAACAAGATTCAGAGGAGAACGCTTTTAAACAGTTTAAATAGATTGCTTTAACAGAA 1273
QY 241 AATTAATATGCAATCTATATTTTGGGAAAAAGATATGATCGGGGTAATATCTTTT 300
| | | | |
Db 1274 AATTAATATGCAATCTATATTTTGGGAAAAAGATATGATCGGGGTAATATCTTTT 1333
QY 301 AACTTTGTCATGAAAAATTAAGTATGAGATGAGATTATCAAGGAATTCAGAAAAATAA 360
```

Db 1334 AACCTTGTCAATGGAATAATGAGTAGATGGAATTATCAACGGAATTCGAGAAATATAA 1393  
Qy 361 ATTGGAGGAATTTATATTTCTTAAGCTCCGAAAGGATGGCAGTAGGAAAAATGGAGCTT 420  
Db 1394 ATTGGAGGAATTTATATTTCTTAAGCTCCGAAAGGATGGCAGTAGGAAAAATGGAGCTT 1453  
Qy 421 ATCAAATGCTGCTCTTCTTCAATCTATTTATTTCCAAAACAAGATGATTTTAAAGAGCTTGG 480  
Db 1454 ATCAATGCTGCTCTTCTTCAATCTATTTATTTCCAAAACAAGATGATTTTAAAGAGCTTGG 1513  
Qy 481 GAAAGAACCAACATGCTAAAGTTTAAATGGAATCATTCAGTAGATGAAAGTAAGTAAG 540  
Db 1514 GAAAGAACCAACATGCTAAAGTTTAAATGGAATCATTCAGTAGATGAAAGTAAGTAAG 1573  
Qy 541 ATTCATTTGAATCCGAAATGGAAGCATTAAGTAGAAGAAATCAATGCTGTGAAGGC 600  
Db 1574 ATTCATTTGAATCCGAAATGGAAGCATTAAGTAGAAGAAATCAATGCTGTGAAGGC 1633  
Qy 601 ATCGGTTTATATGCGCGGATATTAAGATTGAAGAATCTCAATCTAAAGAACAAGAAAT 660  
Db 1634 ATCGGTTTATATGCGCGGATATTAAGATTGAAGAATCTCAATCTAAAGAACAAGAAAT 1693  
Qy 661 ACAGATTTTAAAAATTTAGTCAATATTAAGTATCGAATTAATTTCTGCTGACCGAGAT 720  
Db 1694 ACAGATTTTAAAAATTTAGTCAATATTAAGTATCGAATTAATTTCTGCTGACCGAGAT 1753  
Qy 721 TTAAGAGCTACCAAGCAAAATCTGAGATATTAATTTCTTTCAGCTCACATGATCTCTCT 780  
Db 1754 TTAAGAGCTACCAAGCAAAATCTGAGATATTAATTTCTTTCAGCTCACATGATCTCTCT 1813  
Qy 781 CAAAAAGCTATGGAAAAAAATTCAACTGTGGAAAAAGAAATAGAAATATGTAAAGAA 840  
Db 1814 CAAAAAGCTATGGAAAAAAATTCAACTGTGGAAAAAGAAATAGAAATATGTAAAGAA 1873  
Qy 841 AATACCAAGCAAAATTTGAATCTGATGCTGATTTGGAAGCAGATGAAATATAAAAAAT 900  
Db 1874 AATACCAAGCAAAATTTGAATCTGATGCTGATTTGGAAGCAGATGAAATATAAAAAAT 1933  
Qy 901 AGTGGAAAGCTACAAATGGAGATTTTAAAGAAAGAAAGGAAAAAGAACTTAAATAC 960  
Db 1934 AGTGGAAAGCTACAAATGGAGATTTTAAAGAAAGAAAGGAAAAAGAACTTAAATAC 1993  
Qy 961 ACTCCTTTTAAAGTTTATCAGATGCTGAGAGCTTCCGTAAAGTAATTAAGAAAGTCA 1020  
Db 1994 ACTCCTTTTAAAGTTTATCAGATGCTGAGAGCTTCCGTAAAGTAATTAAGAAAGTCA 2053  
Qy 1021 GGAAGAAATGTTGACATTAAGCTGAGCAAGAAATTTCTATGATGCAACTTTAGTACT 1080  
Db 2054 GGAAGAAATGTTGACATTAAGCTGAGCAAGAAATTTCTATGATGCAACTTTAGTACT 2113  
Qy 1081 AAGCTTGGCAAGACACTTTTAAAGCTTTGTTAAGAGTTCAATTTCTCAATCAATTTAA 1140  
Db 2114 AAGCTTGGCAAGACACTTTTAAAGCTTTGTTAAGAGTTCAATTTCTCAATCAATTTAA 2173  
Qy 1141 GGAATTTTAAAGTTTATGCAAGTAAGTCCAGTGTCTTATTTGAAAGATGCCAAAGTC 1200  
Db 2174 GGAATTTTAAAGTTTATGCAAGTAAGTCCAGTGTCTTATTTGAAAGATGCCAAAGTC 2233  
Qy 1201 GAAAGCAACGAAGAAAGCAAAATATTCATTTCAAGTGAAGTAAGCAACTATGGA 1260  
Db 2234 GAAAGCAACGAAGAAAGCAAAATATTCATTTCAAGTGAAGTAAGCAACTATGGA 2293  
Qy 1261 GAGAGCTACTTCCATTAATAATTTCCAAATTTATTTGGAAGAAAGCCAAATGGAACCT 1320  
Db 2294 GAGAGCTACTTCCATTAATAATTTCCAAATTTATTTGGAAGAAAGCCAAATGGAACCT 2353  
Qy 1321 CTCAGTATCGAGCGGATATATTTCTGCAAAAAAGTAATTCATGTAATCTATTGAAGA 1380  
Db 2354 CTCAGTATCGAGCGGATATATTTCTGCAAAAAAGTAATTCATGTAATCTATTGAAGA 2413  
Qy 1381 GAAATTAATGAAAGGAAGACAGATATTAATCTTCAAAATCTGAAATATCTATTGATGT 1440

Db 2414 GAAATTAATGAAAGGAAGAGAGATATTAATTTCAAAATCTGAAAAATCTATTGATGCT 2473  
Qy 1441 TCTGTTTCTGTGGAAGCAATGAGAGATTCCAATTAAGTAGCTTTTCAGTATTGGTGAAG 1500  
Db 2474 TCTGTTTCTGTGGAAGCAATGAGAGATTCCAATTAAGTAGCTTTTCAGTATTGGTGAAG 2533  
Qy 1501 GAAAGAGAAAAATAATCTTCCGTCAAGATTGCTAAAGAGCAAAAGTAAGATCGAAAA 1560  
Db 2534 GAAAGAGAAAAATAATCTTCCGTCAAGATTGCTAAAGAGCAAAAGTAAGATCGAAAA 2593  
Qy 1561 GATGATGTAATGTAAGAGAGTAAGAGCAATTAATTCATTTGAGAGTCTGTAAAGAGTGA 1620  
Db 2594 GATGATGTAATGTAAGAGAGTAAGAGCAATTAATTCATTTGAGAGTCTGTAAAGAGTGA 2653  
Qy 1621 TTGGGGGATAGTGTAAATGGGGTTGTGGCTGCAATATTTTCTAATCTAATATGCTTCTCC 1680  
Db 2654 TTGGGGGATAGTGTAAATGGGGTTGTGGCTGCAATATTTTCTAATCTAATATGCTTCTCC 2713  
Qy 1681 CGTATAGATGTAATGTAATCTAATGCAATGCAAGAGCACTAAATGTGAGAGCTCATTAAC 1740  
Db 2714 CGTATAGATGTAATGTAATCTAATGCAATGCAAGAGCACTAAATGTGAGAGCTCATTAAC 2773  
Qy 1741 ATTACTAAAAATAGTGTCTGCAAAACAGATCTGATTTGGGAACTTCCAACTTTATGAAT 1800  
Db 2774 ATTACTAAAAATAGTGTCTGCAAAACAGATCTGATTTGGGAACTTCCAACTTTATGAAT 2833  
Qy 1801 GATCAGCTTTATGATCAGGCTCATCTAAATCAATTTTATGATGCAATTAACACAGCGTTT 1860  
Db 2834 GATCAGCTTTATGATCAGGCTCATCTAAATCAATTTTATGATGCAATTAACACAGCGTTT 2893  
Qy 1861 GGAAGAGACAGTGTCAATGAGAAATTAAGATTAAGCTTAACGAATTTATTTAGTGTGCT 1920  
Db 2894 GGAAGAGACAGTGTCAATGAGAAATTAAGATTAAGCTTAACGAATTTATTTAGTGTGCT 2953  
Qy 1921 GTGTCTGCAACCATAGCAAAATCATATTAATTTCTGCTTCTGTGCAATTAAGAGAGTGA 1980  
Db 2954 GTGTCTGCAACCATAGCAAAATCATATTAATTTCTGCTTCTGTGCAATTAAGAGAGTGA 3013  
Qy 1981 AGACTTTCTTCAAGAGTGGAAAGGAGTAAATGAAGGCAATTAAGAACTCAAAATCTT 2040  
Db 3014 AGACTTTCTTCAAGAGTGGAAAGGAGTAAATGAAGGCAATTAAGAACTCAAAATCTT 3073  
Qy 2041 CGAGCGACTACGTCAGTCAAGTGAAGTGTGCTGTACGAAGAAAGAAAGAAAGAACTTAT 2100  
Db 3074 CGAGCGACTACGTCAGTCAAGTGAAGTGTGCTGTACGAAGAAAGAAAGAAAGAACTTAT 3133  
Qy 2101 GGAAGATGAGAGATTTTATTAAGAAATTAATTAATTTCTGCTGCAATTTGCCGAT 2160  
Db 3134 GGAAGATGAGAGATTTTATTAAGAAATTAATTAATTTCTGCTGCAATTTGCCGAT 3193  
Qy 2161 CATGCTGAATTTGATGAGAGAGAAATTAATTAATTAATTTCAAGATGAAATTAATTTGAAT 2220  
Db 3194 CATGCTGAATTTGATGAGAGAGAAATTAATTAATTAATTTCAAGATGAAATTAATTTGAAT 3253  
Qy 2221 AAAAAATCTTCAAAAAATGCAAGTCTGTTATTTGATTAATTAATTTTAAAGAGACT 2280  
Db 3254 AAAAAATCTTCAAAAAATGCAAGTCTGTTATTTGATTAATTTTAAAGAGACT 3313  
Qy 2281 TTTGGAAAAAGAAAGAAAGCTCCAGATATTAATTAATTTGAAATCTATTGAAAAA 2340  
Db 3314 TTTGGAAAAAGAAAGAAAGCTCCAGATATTAATTAATTTGAAATCTATTGAAAAA 3373  
Qy 2341 TTAATTAATGATTTTCAAGAAATTTGATGAAAAACCGAGCTTTTCTAATTAATGATGAA 2400  
Db 3374 TTAATTAATGATTTTCAAGAAATTTGATGAAAAACCGAGCTTTTCTAATTAATGATGAA 3433  
Qy 2401 AGAATGACAAATTAATTTCCGATGAACTTCAAAACAGGAACGCTATTAAGAAATTTGCA 2460  
Db 3434 AGAATGACAAATTAATTTCCGATGAACTTCAAAACAGGAACGCTATTAAGAAATTTGCA 3493  
Qy 2461 AACTATGTTCAAGGAGAAATGAATAATTTAGAGAAAAATTTACGAAAGAAATTTAAAGCT 2520  
Db 3494 AACTATGTTCAAGGAGAAATGAATAATTTAGAGAAAAATTTACGAAAGAAATTTAAAGCT 3553

QY 2521 TTTTCAGAAAGATTGAGTGAAGTGAATTAAGAACTTTGAATTTTACAGAGTAGAAAT 2580  
DB 3554 TTTTCAGAAAGATTGAGTGAAGTGAATTAAGAACTTTGAATTTTACAGAGTAGAAAT 3613  
QY 2581 TATGAAATTTTCACTTTTACCTTCCTCCGAGCTAATGAGAAAGAGTTCCTCT 2640  
DB 3614 TATGAAATTTTCACTTTTACCTTCCTCCGAGCTAATGAGAAAGAGTTCCTCT 3673  
QY 2641 GTGGAGAGAGCTGTTCGTGGGTAGAACGAGAAATTAAGCAAGTATCCGTTGAAAA 2700  
DB 3674 GTGGAGAGAGCTGTTCGTGGGTAGAACGAGAAATTAAGCAAGTATCCGTTGAAAA 3733  
QY 2701 GAGACTAACTTGTCTGCAAAAAAAGATTAAATATPAAAAGCTATCAATPAGCAAAACA 2760  
DB 3734 GAGACTAACTTGTCTGCAAAAAAAGATTAAATATPAAAAGCTATCAATPAGCAAAACA 3793  
QY 2761 GTGAATTTAGTTGGAAATPATTGGAGCTTGCGAGAACAGTACATCCGGAAAGTCCAGTCGGA 2820  
DB 3794 GTGAATTTAGTTGGAAATPATTGGAGCTTGCGAGAACAGTACATCCGGAAAGTCCAGTCGGA 3853  
QY 2821 GGAAGATTAAATGTTCAAAAGATCGAAAAATTCAGCTATGCTAGAAAGCTPAAAGAAAAAGCT 2880  
DB 3854 GGAAGATTAAATGTTCAAAAGATCGAAAAATTCAGCTATGCTAGAAAGCTPAAAGAAAAAGCT 3913  
QY 2881 GAATTAATCAGAGAAATATTTAATGCAATGCAATGCAAGCACTTTTTCATCTAGCCGGA 2940  
DB 3914 GAATTAATCAGAGAAATATTTAATGCAATGCAATGCAAGCACTTTTTCATCTAGCCGGA 3973  
QY 2941 TCTTTTAAATGCTGCTCAGGTGGGAATGCAATCAATGGAATGGGAAGTTATGTGAGGT 3000  
DB 3974 TCTTTTAAATGCTGCTCAGGTGGGAATGCAATCAATGGAATGGGAAGTTATGTGAGGT 4033  
QY 3001 ATCAGTAAAGGCAAGAGTTCCATTTGATGACGAAGCAATATTGAAAAAGCTAATAAAAAAT 3060  
DB 4034 ATCAGTAAAGGCAAGAGTTCCATTTGATGACGAAGCAATATTGAAAAAGCTAATAAAAAAT 4093  
QY 3061 GCTTTAAACAGTAAAGATATCTTCTGTTGGAAATGCTCCGAGTTACCGGGAATCCGA 3120  
DB 4094 GCTTTAAACAGTAAAGATATCTTCTGTTGGAAATGCTCCGAGTTACCGGGAATCCGA 4153  
QY 3121 ACGAAAAATGCGGCGGTCCGAGTTGCTGTTGCGGTAAATGATTAATTTTCAACAA 3180  
DB 4154 ACGAAAAATGCGGCGGTCCGAGTTGCTGTTGCGGTAAATGATTAATTTTCAACAA 4213  
QY 3181 GCTTCATTTGAAGATTAATGACGAAGCAAAAGTAAATATGATPAGAAATAAGATGATGA 3240  
DB 4214 GCTTCATTTGAAGATTAATGACGAAGCAAAAGTAAATATGATPAGAAATAAGATGATGA 4273  
QY 3241 GTPACAGTAATGCGGAATCTTTAGAAATGATGCAAAAAAGCCGGAACAATCAACAGT 3300  
DB 4274 GTPACAGTAATGCGGAATCTTTAGAAATGATGCAAAAAAGCCGGAACAATCAACAGT 4333  
QY 3301 ATTTCTGTTGCGGAGGAATTAATTAAGTTGGAAGTAAACCGAGTGAAGAAAAACCGAA 3360  
DB 4334 ATTTCTGTTGCGGAGGAATTAATTAAGTTGGAAGTAAACCGAGTGAAGAAAAACCGAA 4393  
QY 3361 TCAGAAAGAAAGACCAAGGGAATTTTGGCAAAAATCGGAAACAAATGAGCTCTGTPAAA 3420  
DB 4394 TCAGAAAGAAAGACCAAGGGAATTTTGGCAAAAATCGGAAACAAATGAGCTCTGTPAAA 4453  
QY 3421 AATTAATAATTAACGATAGTATGATTCATTAAACAGAAAAAATTAACAATTTACATTTCTGAA 3480  
DB 4454 AATTAATAATTAACGATAGTATGATTCATTAAACAGAAAAAATTAACAATTTACATTTCTGAA 4513  
QY 3481 GGAATTAATAATTAACGATAGTATGATTCATTAAACAGAAAAAATTAACAATTTACATTTCTGAA 3540  
DB 4514 GGAATTAATAATTAACGATAGTATGATTCATTAAACAGAAAAAATTAACAATTTACATTTCTGAA 4573  
QY 3541 TCTTTCAAGTTGGAGCTTCTGGAAGTGTCTTTCAATTAATTAATAAGAAACATCT 3600  
DB 4574 TCTTTCAAGTTGGAGCTTCTGGAAGTGTCTTTCAATTAATTAATAAGAAACATCT 4633

QY 3601 GCTGCTGAGATGAGTAAAGATTAATTTGAAGGAGCAATTAAGTAGAGTGAAGT 3660  
DB 4634 GCTGCTGAGATGAGTAAAGATTAATTTGAAGGAGCAATTAAGTAGAGTGAAGT 4693  
QY 3661 TCTTCTGATTTCACTTTTGTGTGAGCAATGGGCGGATCTGCTGCACTTCACTGGAATCAT 3720  
DB 4694 TCTTCTGATTTCACTTTTGTGTGAGCAATGGGCGGATCTGCTGCACTTCACTGGAATCAT 4753  
QY 3721 ATTGGAAGTGAATTAAGCAATCACTGCTGCTGTTAGCTGGAAGCGGCTGTGTAATAT 3780  
DB 4754 ATTGGAAGTGAATTAAGCAATCACTGCTGCTGTTAGCTGGAAGCGGCTGTGTAATAT 4813  
QY 3781 ATTCAGATTAACAAGGCTTGGTTGTAATAATGATGATTAATTCGAATGCCAATTAATTT 3840  
DB 4814 ATTCAGATTAACAAGGCTTGGTTGTAATAATGATGATTAATTCGAATGCCAATTAATTT 4873  
QY 3841 AAAAGTAAATGCTTTGAGTGAAGGAATCTCAAGTAGACAGAGACAGGTTTGGAAACAGTT 3900  
DB 4874 AAAAGTAAATGCTTTGAGTGAAGGAATCTCAAGTAGACAGAGACAGGTTTGGAAACAGTT 4933  
QY 3901 AAAGAAAGTGAAGCAAGAAAAAGTTATCTATTGGGAATCTTCTGCTTATCACTTA 3960  
DB 4934 AAAGAAAGTGAAGCAAGAAAAAGTTATCTATTGGGAATCTTCTGCTTATCACTTA 4993  
QY 3961 GTGAACATTAAGTTCGCAAAATCAGAAATTAATCAGTAGACAGAGAAATCTGAAGC 4020  
DB 4994 GTGAACATTAAGTTCGCAAAATCAGAAATTAATCAGTAGACAGAGAAATCTGAAGC 5053  
QY 4021 CAAAAATGGAATGTTGATGATGCTGCTTATCAAGCGGACACCCAAAGTACAGAGACCTTTA 4080  
DB 5054 CAAAAATGGAATGTTGATGATGCTGCTTATCAAGCGGACACCCAAAGTACAGAGACCTTTA 5113  
QY 4081 AATTTCAAGCTGGAAGTCAATATGAACTGTAGGGGCTACTGTGATCTGTTGCCAAATTA 4140  
DB 5114 AATTTCAAGCTGGAAGTCAATATGAACTGTAGGGGCTACTGTGATCTGTTGCCAAATTA 5173  
QY 4141 AAACAACAAGTAAAGCTTCTTATTAAGTGGGAGATTAATCTTAAGCTTAATGAGCGGAC 4200  
DB 5174 AAACAACAAGTAAAGCTTCTTATTAAGTGGGAGATTAATCTTAAGCTTAATGAGCGGAC 5233  
QY 4201 GCAAAAGCTCTTTTGAACAACACTCAAGTGAATGCTGCAAGTGAACGAGGAGCAAT 4260  
DB 5234 GCAAAAGCTCTTTTGAACAACACTCAAGTGAATGCTGCAAGTGAACGAGGAGCAAT 5293  
QY 4261 AGTTCTGAAGCGGATTAAGAAATTAATCAAGGGGCTGTTCTGTCAATTAAGTTGACAA 4320  
DB 5294 AGTTCTGAAGCGGATTAAGAAATTAATCAAGGGGCTGTTCTGTCAATTAAGTTGACAA 5353  
QY 4321 GACGTGAAGCTAGGCTTGAATTAATCTTCCATCGAAGGAGCTAATGAATCAATGTCATT 4380  
DB 5354 GACGTGAAGCTAGGCTTGAATTAATCTTCCATCGAAGGAGCTAATGAATCAATGTCATT 5413  
QY 4381 GCCAAGATGTCAAGGAAGTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGA 4440  
DB 5414 GCCAAGATGTCAAGGAAGTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGA 5473  
QY 4441 AAAAGTAAATAATTTAGAAAGATCTGAGTATTAATGCACTGGAATGCTTAATTAAGC 4500  
DB 5474 AAAAGTAAATAATTTAGAAAGATCTGAGTATTAATGCACTGGAATGCTTAATTAAGC 5533  
QY 4501 AAGGAACAATTAAGAAAAAGCAAGAAAAAGAAAGAGAGGCTAATGTAATGCTGCTTA 4560  
DB 5534 AAGGAACAATTAAGAAAAAGCAAGAAAAAGAAAGAGAGGCTAATGTAATGCTGCTTA 5593  
QY 4561 TCGGTTGCTGGAACGGAATTAATCCGCTGAGAGAGTATGATTAATGCAATCTGTTAAA 4620  
DB 5594 TCGGTTGCTGGAACGGAATTAATCCGCTGAGAGAGTATGATTAATGCAATCTGTTAAA 5653  
QY 4621 AATTAATTTAAAGCAATTAAGTGAAGCAATTAAGAAACCGGAGAGATTAATTCAT 4680  
DB 5654 AATTAATTTAAAGCAATTAAGTGAAGCAATTAAGAAACCGGAGAGATTAATTCAT 5713  
QY 4681 GCGAAACATGTAATGCTGAGGCAAAATCATCTACTGTTGTGTAATGCGGCTCTGGA 4740

D	5714	GCAGAAACATGTAAATGTGAGGCAAAATCATCTAGTGTGTGAAATGCGGCTTCTGGA	5773	D	6794	GCTGAGTTGGAGCTGGAGAGAGCTGAGTGAACAGGTTCTGTAGCAGTGAATTAAGTTGTA	6853
O	4741	CTTGCTATCAGCAAAAGATGCTTTTTCAGGAATGGATCTGGAGCATGGCAAGACTTATCA	4800	O	5821	AATTAATCAGATAGCAGAAATTAATCATGCAAAAATTCATCTGGAGAGGAAATGTGGAGTT	5880
D	5774	CTTGCTATCAGCAAAAGATGCTTTTTCAGGAATGGATCTGGAGCATGGCAAGACTTATCA	5833	D	6854	AATTAATCAGATAGCAGAAATTAATCATGCAAAAATTCATCTGGAGAGGAAATGTGGAGTT	6913
O	4801	AATGACAGCATTCGCAAAAGGTGATTAAGAGAAATTTCTGCTGATTCCTTAATGTGAAC	4860	O	5881	ATTACAGAGTCTGATGCGGTAAATGCTAATTAATGACAGAAACAGTGTCTGAGTGGCCGT	5940
D	5834	AATGACAGCATTCGCAAAAGGTGATTAAGAGAAATTTCTGCTGATTCCTTAATGTGAAC	5893	D	6914	ATTACAGAGTCTGATGCGGTAAATGCTAATTAATGACAGAAACAGTGTCTGAGTGGCCGT	6973
O	4861	GCAAAATTAATTCATTTCTGGGGTGAATGTTGCGGAAACCATTTGCCGTTCTTTCTACG	4920	O	5941	GCAGCATAGGAGCTTCAACCAAGTGTGAATGAATTAACAGGATCTTACAAAAGCATATGTA	6000
D	5894	GCAAAATTAATTCATTTCTGGGGTGAATGTTGCGGAAACCATTTGCCGTTCTTTCTACG	5953	D	6974	GCAGCATAGGAGCTTCAACCAAGTGTGAATGAATTAACAGGATCTTACAAAAGCATATGTA	7033
O	4921	GCGGATAGAGGCTGTTTGGCAATTAATCTCTTCAATTAATAAACCCTGCTTTGATTAACA	4980	O	6001	AAAGATTCCTACAGTGAATGCTAAAGAAAGAAACAGATGTTATTAATTAATCACTCAAGGGCA	6060
D	5954	GCGGATAGAGGCTGTTTGGCAATTAATCTCTTCAATTAATAAACCCTGCTTTGATTAACA	6013	D	7034	AAAGATTCCTACAGTGAATGCTAAAGAAAGAAACAGATGTTATTAATTAATCACTCAAGGGCA	7093
O	4981	GGAACGAAAGTAAATCTTTTGTAGTGAAGAAATACAAAGATCAATGTACAAGCTTTGAAT	5040	O	6061	GTAATTAAGTGTAGATTAAGATTAATTCAAAATCTTAATATTAACGAAGACTTATCAAA	6120
D	6014	GGAACGAAAGTAAATCTTTTGTAGTGAAGAAATACAAAGATCAATGTACAAGCTTTGAAT	6073	D	7094	GTAATTAAGTGTAGATTAAGATTAATTCAAAATCTTAATATTAACGAAGACTTATCAAA	7153
O	5041	GATTTCTCATATTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAAAGCAGCTGGAATC	5100	O	6121	AAAAAGAAATTAAGTAAATTAAGAAAGATTGTTACCAATAGTTCACTCACTATCTTTA	6180
D	6074	GATTTCTCATATTACAAACGTTTCTGCTGAGGCGCTGCAAGTATTAAAGCAGCTGGAATC	6133	D	7154	AAAAAGAAATTAAGTAAATTAAGAAAGATTGTTACCAATAGTTCACTCACTATCTTTA	7213
O	5101	GGAAGAAATGTATCTGTCAAATGCTGTTCTGTAGAAACGGAAGCTTTAGTTAGTAACTCT	5160	O	6181	AAATCTTTATTTGGCAAAATGCGCGTGTTCAGGACAAAGCCGAGTGGCAGGAACGTTAAT	6240
D	6134	GGAAGAAATGTATCTGTCAAATGCTGTTCTGTAGAAACGGAAGCTTTAGTTAGTAACTCT	6193	D	7214	AAATCTTTATTTGGCAAAATGCGCGTGTTCAGGACAAAGCCGAGTGGCAGGAACGTTAAT	7273
O	5161	GAGTTTGAAGAGATTAAGTCTTTCATATGATGCAAAAGATCAAAAACATAAATATACA	5220	O	6241	ATCAACAAGGTTTATGAGAAACAGAAAGCTTTGTAGAAAATCTATATTAATTAATGCAAAA	6300
D	6194	GAGTTTGAAGAGATTAAGTCTTTCATATGATGCAAAAGATCAAAAACATAAATATACA	6253	D	7274	ATCAACAAGGTTTATGAGAAACAGAAAGCTTTGTAGAAAATCTATATTAATTAATGCAAAA	7333
O	5221	ATTGCCGGAATATGCAATGAGAGAAAGCGCTGGAAGTTTGAGCAACAGTTGCTCATACA	5280	O	6301	CATTATTTCTGTAAATCAGAGGATTTACAGCAATTCATCGGAGTATGATGTTCTGTGGT	6360
D	6254	ATTGCCGGAATATGCAATGAGAGAAAGCGCTGGAAGTTTGAGCAACAGTTGCTCATACA	6313	D	7334	CATTATTTCTGTAAATCAGAGGATTTACAGCAATTCATCGGAGTATGATGTTCTGTGGT	7393
O	5281	AATATTGGAAAACATTCAGTTTACTAGCTATTGTAAAAAACAGTAAATTAACAACGCGGAT	5340	O	6361	GTTGGTGGAAATGTATAGGATGAGAGCTTCTCTGATACCAATTAATTAATTAATTAATAC	6420
D	6314	AATATTGGAAAACATTCAGTTTACTAGCTATTGTAAAAAACAGTAAATTAACAACGCGGAT	6373	D	7394	GTTGGTGGAAATGTATAGGATGAGAGCTTCTCTGATACCAATTAATTAATTAATTAATAC	7453
O	5341	GATCAAGATAGAAAAATATCAATGTGACTGCAAAAAGATTATCATATGCAACATATCTATA	5400	O	6421	AAAGCAAGAGTTTGAAGAAAACTTACATGTCTGATGAAGGTTTCCGAGAAAGAACTGAAAT	6480
D	6374	GATCAAGATAGAAAAATATCAATGTGACTGCAAAAAGATTATCATATGCAACATATCTATA	6433	D	7454	AAAGCAAGAGTTTGAAGAAAACTTACATGTCTGATGAAGGTTTCCGAGAAAGAACTGAAAT	7513
O	5401	GCAAGTCGAGTTTGAAGAGCAAAAGAGCTCTGTGCAAGAGCTTTGCAAGTACTAC	5460	O	6481	ACAGCAGATTTTAAGCAAGGAATTTCTCTTTTGAAGTCGAGTCGACGCGCGGTA	6540
D	6434	GCAAGTCGAGTTTGAAGAGCAAAAGAGCTCTGTGCAAGAGCTTTGCAAGTACTAC	6493	D	7514	ACAGCAGATTTTAAGCAAGGAATTTCTCTTTTGAAGTCGAGTCGACGCGCGGTA	7573
O	5461	TTGAATTAAGACAGTTTCTTCTCAATGTTGATCAAACTGATATTAAGAAAGATTAAAGAA	5520	O	6541	GGAAGCCGAGTGGCAGGAAACGTTTCCGTAATCAATTTGCAAGAAAGACGGAAGTGAAT	6600
D	6494	TTGAATTAAGACAGTTTCTTCTCAATGTTGATCAAACTGATATTAAGAAAGATTAAAGAA	6553	D	7574	GGAAGCCGAGTGGCAGGAAACGTTTCCGTAATCAATTTGCAAGAAAGACGGAAGTGAAT	7633
O	5521	GAAATTAATGAAATTAAGGAAAGGCAAAATGTTATGTTCTAGCTGAAATTCAGAGTCAA	5580	O	6601	GTTGAAGAGCAAGAAATTTTGTGTAAAAAAGCTGAGATTAACGCAAAAACGTTATAGTTCT	6660
D	6554	GAAATTAATGAAATTAAGGAAAGGCAAAATGTTATGTTCTAGCTGAAATTCAGAGTCAA	6613	D	7634	GTTGAAGAGCAAGAAATTTTGTGTAAAAAAGCTGAGATTAACGCAAAAACGTTATAGTTCT	7693
O	5581	GTTGTCACAAATGCAACAGTGTCTTTCCGAGCAAGTGAACAAAGCTGCAAGTGAAGCTGGA	5640	O	6661	GTTGCAATTTGGAATTTGCCGAGCTGCGAGTGTGCTGCAAAAAGAGCTGCAATTTGAGACACA	6720
D	6614	GTTGTCACAAATGCAACAGTGTCTTTCCGAGCAAGTGAACAAAGCTGCAAGTGAAGCTGGA	6673	D	7694	GTTGCAATTTGGAATTTGCCGAGCTGCGAGTGTGCTGCAAAAAGAGCTGCAATTTGAGACACA	7753
O	5641	GTAAGAGTTTAATTAATAATTAACAATAATCTTCTGCAATATAAATAATTAATTAATTAAT	5700	O	6721	GTTGCAATTTGGAATTTGCCGAGCTGCGAGTGTGCTGCAAAAAGAGCTGCAATTTGAGACACA	6780
D	6674	GTAAGAGTTTAATTAATAATTAACAATAATCTTCTGCAATATAAATAATTAATTAATTAAT	6733	D	7754	GTTGCAATTTGGAATTTGCCGAGCTGCGAGTGTGCTGCAAAAAGAGCTGCAATTTGAGACACA	7813
O	5701	GTAAGAGTTTAATTAATAATTAACAATAATCTTCTGCAATATAAATAATTAATTAATTAAT	5760	D	6781	ACTCGAAACCAAGTTTGAATGATATAGCAGAAATAGATTAATTAATTAATTAATTAATTAAT	6840
D	6734	GTAAGAGTTTAATTAATAATTAACAATAATCTTCTGCAATATAAATAATTAATTAATTAAT	6793	D	7814	ACTCGAAACCAAGTTTGAATGATATAGCAGAAATAGATTAATTAATTAATTAATTAATTAAT	7873
O	5761	GCTGAGTTTGAAGCTGAGAGAGCTGAGTGAACAGGTTCTGTAGCAGTGAATTAAGATTGTA	5820	O	6841	TCAGCCGAGAGCTGGAATTTCTTTCAGCCGAGAGTATCTGAGAGTGTGTTCTGTCAATTAATTT	6900
D				D	7874	TCAGCCGAGAGCTGGAATTTCTTTCAGCCGAGAGTATCTGAGAGTGTGTTCTGTCAATTAATTT	7933

QY 6901 GCAAAATAGGTGAAACAGATATCGATCATAGTACTTTACACTCTTTCTTACTGATTAAT 6960  
|  
|  
|  
Db 7934 GCAATTAAGGTGAAACAGATATCGATCATAGTACTTTACACTCTTTCTTACTGATTAAT 7993  
|  
QY 6961 GTAAAAGCTTTAATTAATTTGAAATTCCTTGAACGCCGGTGGAGAGCCGACAGTCTT 7020  
|  
|  
|  
Db 7994 GTAAAAGCTTTAATTAATTTGAAATTCCTTGAACGCCGGTGGAGAGCCGACAGTCTT 8053  
|  
QY 7021 GCACAGTTTACCCGAGTGGTTCCTGTTAACATTAATTAAGTTCTGTGATAGCTGAGTT 7080  
|  
|  
|  
Db 8054 GCAAGAGTTTACCCGAGTGGTTCCTGTTAACATTAATTAAGTTCTGTGATAGCTGAGTT 8113  
|  
QY 7081 CACAAATACCTGATTTGACTTCCTGTAGAGAAAAAGTAATGTAAACGCCCAAAAGAGAA 7140  
|  
|  
|  
Db 8114 CACAAATACCTGATTTGACTTCCTGTAGAGAAAAAGTAATGTAAACGCCCAAAAGAGAA 8173  
|  
QY 7141 AAAAATATTGAACAAACGACGCAAAATGACGAAATCCGAGGAGCAATCGAGCCAAAT 7200  
|  
|  
|  
Db 8174 AAAAATATTGAACAAACGACGCAAAATGACGAAATCCGAGGAGCAATCGAGCCAAAT 8233  
|  
QY 7201 GTCTTGTAATTAATTTTGGAAACAGCTGTAGAAAGTAGAAAAAATTTGAAAGAAAGGA 7260  
|  
|  
|  
Db 8234 GTCTTGTAATTAATTTTGGAAACAGCTGTAGAAAGTAGAAAAAATTTGAAAGAAAGGA 8293  
|  
QY 7261 ACAGAAAGTTTAAAAAATTGACAGAGTTAAACAAAGAACAGATTAATAATGAT 7320  
|  
|  
|  
Db 8294 ACAGAAAGTTTAAAAAATTGACAGAGTTAAACAAAGAACAGATTAATAATGAT 8353  
|  
QY 7321 GCTAGAAAAAATCTTCAATCAGAGGTAATTTCTACAGAGATTAATCTGTAAGAGC 7380  
|  
|  
|  
Db 8354 GCTAGAAAAAATCTTCAATCAGAGGTAATTTCTACAGAGATTAATCTGTAAGAGC 8413  
|  
QY 7381 GATAGAGGATATCTCAGGAGAGAAATTAAGCCATTGTGAGAGCTTCTGATTTAT 7440  
|  
|  
|  
Db 8414 GATAGAGGATATCTCAGGAGAGAAATTAAGCCATTGTGAGAGCTTCTGATTTAT 8473  
|  
QY 7441 GAAAAAATATGATATTAACAACAGAGCAAGATTAATATCATCTTCTGAGTGGTTG 7500  
|  
|  
|  
Db 8474 GAAAAAATATGATATTAACAACAGAGCAAGATTAATATCATCTTCTGAGTGGTTG 8533  
|  
QY 7501 GGAATCGCAGGCTTGTCCGATCAGAGAACAGTGGCAGTTAAATTTAAAAAGAAAT 7560  
|  
|  
|  
Db 8534 GGAATCGCAGGCTTGTCCGATCAGAGAACAGTGGCAGTTAAATTTAAAAAGAAAT 8593  
|  
QY 7561 TCCGAGTTACTGTGAAAAATCTTTGTGAAAGCAGCTGAAAAAGTAATGTAGATCG 7620  
|  
|  
|  
Db 8594 TCCGAGTTACTGTGAAAAATCTTTGTGAAAGCAGCTGAAAAAGTAATGTAGATCG 8653  
|  
QY 7621 GATATTTACAGAAATGTTGCTTTAACAGCAATCAAGGTCTGTAGAGCAATTTGGAAAT 7680  
|  
|  
|  
Db 8654 GATATTTACAGAAATGTTGCTTTAACAGCAATCAAGGTCTGTAGAGCAATTTGGAAAT 8713  
|  
QY 7681 GGAGCTGCTTAATGCAAAATTAATTTCTAATGGAAGTCAAAATATCATCTTAAAAATTT 7740  
|  
|  
|  
Db 8714 GGAGCTGCTTAATGCAAAATTAATTTCTAATGGAAGTCAAAATATCATCTTAAAAATTT 8773  
|  
QY 7741 AAGCTATTAGAAAAAATATGATGTTATTGTAAAAAGATTAATCGAAATTTGAGCGGAA 7800  
|  
|  
|  
Db 8774 AAGCTATTAGAAAAAATATGATGTTATTGTAAAAAGATTAATCGAAATTTGAGCGGAA 8833  
|  
QY 7801 GCAAAAAGATTAACCGTAGAGAGCGGTAGCTGCCGAGCCATTAATCTCAAAAAGCAAGAAAT 7860  
|  
|  
|  
Db 8834 GCAAAAAGATTAACCGTAGAGAGCGGTAGCTGCCGAGCCATTAATCTCAAAAAGCAAGAAAT 8893  
|  
QY 7861 GAAATGAATTCAGAGGTTGAAATTTGAGAGAGATTTTCAATGAAGAAATAGAGTAAT 7920  
|  
|  
|  
Db 8894 GAAATGAATTCAGAGGTTGAAATTTGAGAGAGATTTTCAATGAAGAAATAGAGTAAT 8953  
|  
QY 7921 AGCCCTTTCAAAGAAATTCGAAAGAGAAATCAATGTCAAAGTGAAAAAGAAAAACAAGTGC 7980  
|  
|  
|  
Db 8954 AGCCCTTTCAAAGAAATTCGAAAGAGAAATCAATGTCAAAGTGAAAAAGAAAAACAAGTGC 9013

QY 7981 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGTACGAGGGGACAGAAATTAATTTCCGAA 8040  
|  
|  
|  
Db 9014 ACTGCTGAATCTCAAGAGCTTCTGTAGAGACAGTACGAGGGGACAGAAATTAATTTCCGAA 9073  
|  
QY 8041 GCAAAAAGTCCGGAAGCTCTTAATTTGAAAGTTAGTACAAAATCCGAAAGAAATTTTTT 8100  
|  
|  
|  
Db 9074 GCAAAAAGTCCGGAAGCTCTTAATTTGAAAGTTAGTACAAAATCCGAAAGAAATTTTTT 9133  
|  
QY 8101 CATGCAATTAATGTGAAATTAATGGAAGCAACATTAATAAGAAAGTAAACGAGTTTCTAAA 8160  
|  
|  
|  
Db 9134 CATGCAATTAATGTGAAATTAATGGAAGCAACATTAATAAGAAAGTAAACGAGTTTCTAAA 9193  
|  
QY 8161 GCAGTAAACAGGTTCTGTATTGGAGAGAGTGAAGTCAACCAAGCGAAGTAACTACTGCA 8220  
|  
|  
|  
Db 9194 GCAGTAAACAGGTTCTGTATTGGAGAGAGTGAAGTCAACCAAGCGAAGTAACTACTGCA 9253  
|  
QY 8221 GGTAAACCTATGTGTAAGAGTTGAGGAAGAAATTTGTTCAAGAACAAATCGATTGAATGCA 8280  
|  
|  
|  
Db 9254 GGTAAACCTATGTGTAAGAGTTGAGGAAGAAATTTGTTCAAGAACAAATCGATTGAATGCA 9313  
|  
QY 8281 ATTTCTAAAGTGAAGGTTTGGATGCAAGTAAAGTAACTGCTTAAATCTTCTGTAGTATCA 8340  
|  
|  
|  
Db 9314 ATTTCTAAAGTGAAGGTTTGGATGCAAGTAAAGTAAAGTAACTGCTTAAATCTTCTGTAGTATCA 9373  
|  
QY 8341 GGAATAATGAGAGGAATTTGCCGAGCAGAGTGAATTACTTCTACAGACAAAGTAATACT 8400  
|  
|  
|  
Db 9374 GGAATAATGAGAGGAATTTGCCGAGCAGAGTGAATTACTTCTACAGACAAAGTAATACT 9433  
|  
QY 8401 GAATCCGTAGTTCGTTTACGAAAGCAGATTAATGAATAATGATTACACAAAAAATAT 8460  
|  
|  
|  
Db 9434 GAATCCGTAGTTCGTTTACGAAAGCAGATTAATGAATAATGATTACACAAAAAATAT 9493  
|  
QY 8461 ATTTGAGAGTCAATGCTCTTCTTTAAATGATACAAAGAAATGAAGCCGAAATTAATCT 8520  
|  
|  
|  
Db 9494 ATTTGAGAGTCAATGCTCTTCTTTAAATGATACAAAGAAATGAAGCCGAAATTAATCT 9553  
|  
QY 8521 TTAGCGGTAGCCGGTGCATGACACAGGAACAAACAAAGCAATTAACGATCAAAACAAG 8580  
|  
|  
|  
Db 9554 TTAGCGGTAGCCGGTGCATGACACAGGAACAAACAAAGCAATTAACGATCAAAACAAG 9613  
|  
QY 8581 TTAACTTCTACAACTGTAAATGAGAGAAACGTATCTCAACTTCTGTCAAAAACCTTTGCT 8640  
|  
|  
|  
Db 9614 TTAACTTCTACAACTGTAAATGAGAGAAACGTATCTCAACTTCTGTCAAAAACCTTTGCT 9673  
|  
QY 8641 AAAAAATGAATTAATGAAATGTAAGAACTGAGAGAGCTTGTGAGACCGGAAACA 8700  
|  
|  
|  
Db 9674 AAAAAATGAATTAATGAAATGTAAGAACTGAGAGAGCTTGTGAGACCGGAAACA 9733  
|  
QY 8701 GCAGCCGTTGAAAAATTAATCAAGAGTACTACAGAGCAATTTGTCAGAAATTTGGGAA 8760  
|  
|  
|  
Db 9734 GCAGCCGTTGAAAAATTAATCAAGAGTACTACAGAGCAATTTGTCAGAAATTTGGGAA 9793  
|  
QY 8761 ATTTGAGATTAATTAAGAAACGATTCAGAGATTAATCAATTTGTAAGATCAACGAGAC 8820  
|  
|  
|  
Db 9794 ATTTGAGATTAATTAAGAAACGATTCAGAGATTAATCAATTTGTAAGATCAACGAGAC 9853  
|  
QY 8821 GGAACCAAGAGGCTTGTCCGAAAGAAATGTAATTTCTGTGAAAAATCAATTTCAAGG 8880  
|  
|  
|  
Db 9854 GGAACCAAGAGGCTTGTCCGAAAGAAATGTAATTTCTGTGAAAAATCAATTTCAAGG 9913  
|  
QY 8881 GAAACAAAATCAATTCATGGAAGATTAAGCAGAAATTTGTTGAAACCGGAAGTAAATGTA 8940  
|  
|  
|  
Db 9914 GAAACAAAATCAATTCATGGAAGATTAAGCAGAAATTTGTTGAAACCGGAAGTAAATGTA 9973  
|  
QY 8941 GATGCTTTGAATGAACCTGATGTAATCTAACAGGAAAAAGTGGTCTATGCTGGAATTT 9000  
|  
|  
|  
Db 9974 GATGCTTTGAATGAACCTGATGTAATCTAACAGGAAAAAGTGGTCTATGCTGGAATTT 10033  
|  
QY 9001 GGTATTTGAAATGTTGATGTAATTAATGATGTTAAAGAAAAATGTAACCCAAAATCCGA 9060  
|  
|  
|  
Db 10034 GGTATTTGAAATGTTGATGTAATTAATGATGTTAAAGAAAAATGTAACCCAAAATCCGA 10093  
|  
QY 9061 AGACATGCTAATTTGTGAAGCTACTGGAAGAAACAAGAAATATCAAGCAATTTCAAGAGCAAAA 9120

Db 10094 AGACATGCTATGTGAGAACTCTGAGAAACAGAAATATCAAGCATTTACAGAGACAAAA 10153  
Qy 9121 GTAATATTTCTGGAGAAAGAGACCGCTGACGTGACGTGCATATTCGAATGTACACTT 9180  
Db 10154 GTAATATTTCTGGAGAAAGAGACCGCTGACGTGACGTGCATATTCGAATGTACACTT 10213  
Qy 9181 TCCATGAGATGATATTTAAAAATTTGGCAAGCAGTATGCATCTTCTCAATTAATAACC 9240  
Db 10214 TCCATGAGATGATATTTAAAAATTTGGCAAGCAGTATGCATCTTCTCAATTAATAACC 10273  
Qy 9241 AAAAAATCAAAAAATTAATTTACTTTAGCATCAAGTGAATGCAATGTGAATGTTCAAT 9300  
Db 10274 AAAAAATCAAAAAATTAATTTACTTTAGCATCAAGTGAATGCAATGTGAATGTTCAAT 10333  
Qy 9301 GGGGCTGAG 9360  
Db 10334 GGGGCTGAG 10393  
Qy 9361 AATAGAACTAATATATGTTGATTTAGCAGAGAAATTTAAAAACAGAGAGAAACATCAATGTA 9420  
Db 10394 AATAGAACTAATATATGTTGATTTAGCAGAGAAATTTAAAAACAGAGAGAAACATCAATGTA 10453  
Qy 9421 TATGCCGATATGATPAAAAATTAATATATAGTAAGACAAATTTCTAAGGCTATTGCGGAT 9480  
Db 10454 TATGCCGATATGATPAAAAATTAATATATAGTAAGACAAATTTCTAAGGCTATTGCGGAT 10513  
Qy 9481 GCCAAAGCTATGCTGCGAGCTGCTCGGCAACTGCGCATATGAGAAAAATGAAATGATAAA 9540  
Db 10514 GCCAAAGCTATGCTGCGAGCTGCTCGGCAACTGCGCATATGAGAAAAATGAAATGATAAA 10573  
Qy 9541 TTTAATATATGCGATCCGAGAAATTTTAAAAATATCTGCAAGATTTGGAAGGAAAGCTAAT 9600  
Db 10574 TTTAATATATGCGATCCGAGAAATTTTAAAAATATCTGCAAGATTTGGAAGGAAAGCTAAT 10633  
Qy 9601 AAAAAACGTGCGATGATCTTATCAGGTAGCTGTATACGATTAATATATCATGTCAT 9660  
Db 10634 AAAAAACGTGCGATGATCTTATCAGGTAGCTGTATACGATTAATATATCATGTCAT 10693  
Qy 9661 TCTTCTGAAAAAGCATACAAAAAATTTGACATTCATCAATCAAGAGAGAGAGAGAGAGAG 9720  
Db 10694 TCTTCTGAAAAAGCATACAAAAAATTTGACATTCATCAATCAAGAGAGAGAGAGAGAGAG 10753  
Qy 9721 AAATGA 9726  
Db 10754 AAATGA 10759

RESULT 5  
US-09-841-786-10  
; Sequence 10, Application US/09841786  
; Patent No. US20020054883A1  
; GENERAL INFORMATION:  
; APPLICANT: NAGARAJA, T. G.  
; APPLICANT: STEWART, GEORGE C.  
; APPLICANT: NARAYANAN, SANJEEV K.  
; APPLICANT: CHENGAPPA, M. M.  
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF  
; FILE REFERENCE: 30296  
; CURRENT APPLICATION NUMBER: US/09/841,786  
; CURRENT FILING DATE: 2001-04-24  
; PRIOR APPLICATION NUMBER: 09/558,257  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 2780  
; TYPE: DNA  
; ORGANISM: Fusobacterium necrophorum  
US-09-841-786-10

Query Match 28.6%; Score 2780; DB 3; Length 2780;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 919 GGGAGATTTATTAAG 978  
Db 1 GGGAGATTTATTAAG 60  
Qy 979 GATGTGAGAGGCTCCGTAAAGATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1038  
Db 61 GATGTGAGAGGCTCCGTAAAGATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
Qy 1039 ACAGCTGAAGCAAGAAATTTCTATATGATGCAATTTAGTTACTTAAGCTTGCAAGACCTCT 1098  
Db 121 ACAGCTGAAGCAAGAAATTTCTATATGATGCAATTTAGTTACTTAAGCTTGCAAGACCTCT 180  
Qy 1099 TTTAGCTTTGTTACAGGTTCTATTTCTCTATTCATATTAATGAGATTTTAAAGTTTAAAGTTTAA 1158  
Db 181 TTTAGCTTTGTTACAGGTTCTATTTCTCTATTCATATTAATGAGATTTTAAAGTTTAAAGTTTAA 240  
Qy 1159 ACAAGTAAGTCAGTGTGTTATTTGAG 1218  
Db 241 ACAAGTAAGTCAGTGTGTTATTTGAG 300  
Qy 1219 GCAATATTCATTTCTTACAGTGAAG 1278  
Db 301 GCAATATTCATTTCTTACAGTGAAG 360  
Qy 1279 AAAATTAACATTTTATTTTGAAG 1338  
Db 361 AAAATTAACATTTTATTTTGAAG 420  
Qy 1339 TATATTTCTGCAAAAAGATTAATTCATATGTAATTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1398  
Db 421 TATATTTCTGCAAAAAGATTAATTCATATGTAATTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
Qy 1399 AGAGCAAGATTAATTTCTTCAAAATCTGAAAATCTATATGATGCTTCTGTTTCTGTTGAGAGAG 1458  
Db 481 AGAGCAAGATTAATTTCTTCAAAATCTGAAAATCTATATGATGCTTCTGTTTCTGTTGAGAGAG 540  
Qy 1459 ATGAGAGATTTCCATTAAG 1518  
Db 541 ATGAGAGATTTCCATTAAG 600  
Qy 1519 TCCGTCAAGATTTGCTAAAG 1578  
Db 601 TCCGTCAAGATTTGCTAAAG 660  
Qy 1579 AGTGAAGGATTAATTTCCATTCGAGCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1638  
Db 661 AGTGAAGGATTAATTTCCATTCGAGCTGCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Qy 1639 GGGGTTGTGCTGCAAAATATTTCTTCAATATATGCTTCTCCGATATGATGATGATGAGAGAGAGAG 1698  
Db 721 GGGGTTGTGCTGCAAAATATTTCTTCAATATATGCTTCTCCGATATGATGATGATGAGAGAGAGAG 780  
Qy 1699 TATCTACATGCGCAAG 1758  
Db 781 TATCTACATGCGCAAG 840  
Qy 1759 CTGCAACAGGATCTGATTTTGGAG 1818  
Db 841 CTGCAACAGGATCTGATTTTGGAG 900  
Qy 1819 GGTCTATCTAAATCAATTTTATGATGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1878  
Db 901 GGTCTATCTAAATCAATTTTATGATGCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
Qy 1879 GAGGAATTAAGAGATTAAG 1938  
Db 961 GAGGAATTAAGAGATTAAG 1020  
Qy 1939 AATCATATTAATTTCTGCTTCTGTGCAATAG 1998

Db 1021 AATCATATAATTCCTGCTTCTGTGSCAATAGAGAGATGGAAGACTTTCTTCAGAGATG 1080  
Qy 1999 GAAGGAGCTAAATGTAAAGGCACTTAAATGAAGCTCAAAATCTTCGAGCACTACGTAAGT 2058  
Db 1081 GAAGGAGCTAAATGTAAAGGCACTTAAATGAAGCTCAAAATCTTCGAGCACTACGTAAGT 1140  
Qy 2059 GGAAGTGTGCTGTGTCGAAAG 2118  
Db 1141 GGAAGTGTGCTGTGTCGAAAG 1200  
Qy 2119 TATGGAATCTATTAATAATATATCTGTGCAATTCGCGATCATGCTGAATGGTATCG 2178  
Db 1201 TATGGAATCTATTAATAATATATCTGTGCAATTCGCGATCATGCTGAATGGTATCG 1260  
Qy 2179 GAAGGAG 2238  
Db 1261 GAAGGAG 1320  
Qy 2239 GCAAGTCTGTATGATTAATTAAGACTTTTAAAGAGAGCTTTTGAAGAGAGAGAGAGAGAG 2298  
Db 1321 GCAAGTCTGTATGATTAATTAAGACTTTTAAAGAGAGCTTTTGAAGAGAGAGAGAGAGAG 1380  
Qy 2299 ACTCCAGATATGATCCGAAAGATATGTAATCTATTTGAGAGAGAGAGAGAGAGAGAGAGAG 2358  
Db 1381 ACTCCAGATATGATCCGAAAGATATGTAATCTATTTGAGAGAGAGAGAGAGAGAGAGAGAG 1440  
Qy 2359 GAAAAATTTGATGAG 2418  
Db 1441 GAAAAATTTGATGAG 1500  
Qy 2419 CCGGATGGAATCTCAAAAG 2478  
Db 1501 CCGGATGGAATCTCAAAAG 1560  
Qy 2479 ATGAAGAGATTTAG 2538  
Db 1561 ATGAAGAGATTTAG 1620  
Qy 2539 GAGAGATTTAAAG 2598  
Db 1621 GAGAGATTTAAAG 1680  
Qy 2599 TTTTACCTCTTCGAGAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2658  
Db 1681 TTTTACCTCTTCGAGAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
Qy 2659 TGGGTAG 2718  
Db 1741 TGGGTAG 1800  
Qy 2719 AAAAAAGATTTAAATTAATAAGCTATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2778  
Db 1801 AAAAAAGATTTAAATTAATAAGCTATCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
Qy 2779 ATTGACCTTGCAG 2838  
Db 1861 ATTGACCTTGCAG 1920  
Qy 2839 AAGATGAG 2898  
Db 1921 AAGATGAG 1980  
Qy 2899 ATTAATGAG 2958  
Db 1981 ATTAATGAG 2040  
Qy 2959 GGTGGAGATGCAATCAATGAG 3018  
Db 2041 GGTGGAGATGCAATCAATGAG 2100  
Qy 3019 TCCATTTGAG 3078  
Db 2101 TCCATTTGAG 2160

Qy 3079 GATACCTTCTGTGGAATGCTGCCGCTTCACGCGGAAATCGAAGCAAAATGCGCGGCTC 3138  
Db 2161 GATACCTTCTGTGGAATGCTGCCGCTTCACGCGGAAATCGAAGCAAAATGCGCGGCTC 2220  
Qy 3139 GGGGTGCTGTGTGGGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3198  
Db 2221 GGGGTGCTGTGTGGGATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280  
Qy 3199 GACGAGAGCAAAAGTAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3258  
Db 2281 GACGAGAGCAAAAGTAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
Qy 3259 TCTTTAGAGATGATGCAAAAG 3318  
Db 2341 TCTTTAGAGATGATGCAAAAG 2400  
Qy 3319 ATTAATAAGTGTGGAAGTAAACCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3378  
Db 2401 ATTAATAAGTGTGGAAGTAAACCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2460  
Qy 3379 GGAATTTTGGCAAAATCGAAGCAAAAGTGAATCTGTGTAATAATAATAATAATAATAATAATAAT 3438  
Db 2461 GGAATTTTGGCAAAATCGAAGCAAAAGTGAATCTGTGTAATAATAATAATAATAATAATAATAAT 2520  
Qy 3439 ATGATTTATTAACAG 3498  
Db 2521 ATGATTTATTAACAG 2580  
Qy 3499 AATCTTCTTGAAG 3558  
Db 2581 AATCTTCTTGAAG 2640  
Qy 3559 TCTGGAAGTGTTCCTTCAATTAATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3618  
Db 2641 TCTGGAAGTGTTCCTTCAATTAATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Qy 3619 AAGATTAATTTGAAG 3678  
Db 2701 AAGATTAATTTGAAG 2760  
Qy 3679 GTTGAGAGATGGGCGGATC 3698  
Db 2761 GTTGAGAGATGGGCGGATC 2780

RESULT 6  
US-10-647-057-10  
; Sequence 10, Application US/10647057  
; Publication No. US20040047871A1  
; GENERAL INFORMATION:  
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION  
; APPLICANT: NAGARAJA, T.  
; APPLICANT: STEWART, GEORGE  
; APPLICANT: NARAYANAN, SANJEEV  
; APPLICANT: CHENGAPPA, M.  
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP  
; FILE REFERENCE: PCT-30962  
; CURRENT APPLICATION NUMBER: US/10/647,057  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patencin version 3.0  
; SEQ ID NO 10  
; LENGTH: 2780  
; TYPE: DNA  
; ORGANISM: Fusobacterium necrophorum  
US-10-647-057-10

Query Match 28.6%; Score 2780; DB 7; Length 2780;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY GGGAGATTTATTAAGAAAAGAGGAGGAGAACTTATACACTCTTTAAAGTTATCA 978  
Db GGGAAATTTATTAAGAAAAGAGGAGGAGAACTTATACACTCTTTAAAGTTATCA 60  
QY GATGTGGAAGCTTCGGTAAAGCTTAATAAGAAAAGTCATGGAAGAAAGTTTACATT 1038  
Db GATGTGGAAGCTTCGGTAAAGCTTAATAAGAAAAGTCATGGAAGAAAGTTTACATT 120  
QY ACAGCTGAAGCAAAAGATTTCTATGATGCAACTTATAGTACTAGCTTGCAAGCACTCT 1098  
Db ACAGCTGAAGCAAAAGATTTCTATGATGCAACTTATAGTACTAGCTTGCAAGCACTCT 180  
QY TTTACCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAATGATTTTATAGTTTATG 1158  
Db TTTACCTTTGTTACAGGTTCTATTTCTCTATCAATTTAAATGATTTTATAGTTTATG 240  
QY ACAATTAAGTCCAGTGTGCTTATTTGAAAAGATGCCAAGTGAAGCAACAGAGAAAG 1218  
Db ACAGTAAAGTCCAGTGTGCTTATTTGAAAAGATGCCAAGTGAAGCAACAGAGAAAG 300  
QY GCAATTTATTCATTTCTTACAGTGAAGCACTATGGAAGCACTTCTTCATTTA 1278  
Db GCAATTTATTCATTTCTTACAGTGAAGCACTATGGAAGCACTTCTTCATTTA 360  
QY AAAATTTACCAATTTATTTGGAAGCAATGCAAACTTCTCAGTATCGAGCGGGA 1338  
Db AAAATTTACCAATTTATTTGGAAGCAATGCAAACTTCTCAGTATCGAGCGGGA 420  
QY TATATTTCTGCAAAAAGTAACTTCCAAATGTAATTTGAAGGAAGTAAATTCGAAGGA 1398  
Db TATATTTCTGCAAAAAGTAACTTCCAAATGTAATTTGAAGGAAGTAAATTCGAAGGA 480  
QY AGAGCAGATATTAATCTCAAAATCTGAAAATCTATTTAGTCTTCTGTTTCTGTTGAAG 1458  
Db AGAGCAGATATTAATCTCAAAATCTGAAAATCTATTTAGTCTTCTGTTTCTGTTGAAG 540  
QY ATGAGAGATTCCAATTAAGTAGCTCTTTCAGTATTTGTTGGAAGCAAGAAAATTAATCT 1518  
Db ATGAGAGATTCCAATTAAGTAGCTCTTTCAGTATTTGTTGGAAGCAAGAAAATTAATCT 600  
QY TCCGTCAGAGTTGCTTAAAGAGCAAAAGTAAATGCAAAACGAGATGTAATTTGGA 1578  
Db TCCGTCAGAGTTGCTTAAAGAGCAAAAGTAAATGCAAAACGAGATGTAATTTGGA 660  
QY AGTGAAGCGATTAATTCATTCGAGCTGCTGTAAGAAGTGAAGTGGGGATAGTGTAT 1638  
Db AGTGAAGCGATTAATTCATTCGAGCTGCTGTAAGAAGTGAAGTGGGGATAGTGTAT 720  
QY GGGGTTGTGCTGCAAAATATTTCTAATAATGCTTCTCCGTAATGATGTAGTGA 1698  
Db GGGGTTGTGCTGCAAAATATTTCTAATAATGCTTCTCCGTAATGATGTAGTGA 780  
QY TATCTAATGCCAAGAGCACTAAATGTGAGGCTCATTAATTAATTAATAGTGT 1758  
Db TATCTAATGCCAAGAGCACTAAATGTGAGGCTCATTAATTAATTAATAGTGT 840  
QY CTGCAACAAGATCTGATTTGGGAAGCTTCAAGTTATGATGATCAGGTTTATATCA 1818  
Db CTGCAACAAGATCTGATTTGGGAAGCTTCAAGTTATGATGATCAGGTTTATATCA 900  
QY GGTCTATCTTAAATCAATTTTATAGTCAATTAACAAGCGGTTTGAAGAGCAGTGTCAAT 1878  
Db GGTCTATCTTAAATCAATTTTATAGTCAATTAACAAGCGGTTTGAAGAGCAGTGTCAAT 960  
QY GAGGAAATTAAGAAATTAAGCTTAAGCTTATAGTGTGCTGCAACCTATGA 1938  
Db GAGGAAATTAAGAAATTAAGCTTAAGCTTATAGTGTGCTGCAACCTATGA 1020  
QY AATCATTAATTAATTTGCTTCTGCTGCAATAGAGAGAGTGAAGCTTCTTCAAGAGTG 1998  
Db AATCATTAATTAATTTGCTTCTGCTGCAATAGAGAGAGTGAAGCTTCTTCAAGAGTG 1080  
QY GAAAGGAGTAAATGTAAGGCAATTAATGAAGCTCAAAATCTTGAAGCACTACGTCAGT 2058

Db GAAAGGAGTAAATGTAAGGCAATTAATGAAGCTCAAAATCTTGAAGCACTACGTCAGT 1140  
QY GGAAGTGTGCTGTATCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2118  
Db GGAAGTGTGCTGTATCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
QY TATGGAATCTTAAATTAATAGTCTTCTGTGCAATTTGCGCATGCTGAATTTGGTATCG 2178  
Db TATGGAATCTTAAATTAATAGTCTTCTGTGCAATTTGCGCATGCTGAATTTGGTATCG 1260  
QY GAAGAGAAATTTGATATCAACAGTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 2238  
Db GAAGAGAAATTTGATATCAACAGTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1320  
QY GCAAGCTGTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2298  
Db GCAAGCTGTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380  
QY ACTCCAGATATGATCCGAAGAGATTTGAATCTATTTGAATTAATTAATTAATTAATTAAT 2358  
Db ACTCCAGATATGATCCGAAGAGATTTGAATCTATTTGAATTAATTAATTAATTAATTAAT 1440  
QY GAAAAATTTGATGAGAAAAACCGAGCTTTTACTAATGATGAGAAATGACAAATTAATCTT 2418  
Db GAAAAATTTGATGAGAAAAACCGAGCTTTTACTAATGATGAGAAATGACAAATTAATCTT 1500  
QY CCGATGGAATTTGAAAAAGCAGAACTGCTATTAAGAAATTTGAACTATGTTCAAGGAGAA 2478  
Db CCGATGGAATTTGAAAAAGCAGAACTGCTATTAAGAAATTTGAACTATGTTCAAGGAGAA 1560  
QY ATGAAAAAATTTAGAGAGAAAAATTAACGAAAGATTTAAAGCTTTTTCAGAGAGATTTGAGT 2538  
Db ATGAAAAAATTTAGAGAGAAAAATTAACGAAAGATTTAAAGCTTTTTCAGAGAGATTTGAGT 1620  
QY GGAATGATTAAGAAACCTTTGAATTTTACAGAGTGAAGAAATTAATGCAATTTTCAACT 2598  
Db GGAATGATTAAGAAACCTTTGAATTTTACAGAGTGAAGAAATTAATGCAATTTTCAACT 1680  
QY TTTTACCTTTCCGAGAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2658  
Db TTTTACCTTTCCGAGAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
QY TGGGTAGAACAGAGAAATTTATAGCAAGTATCCGTTGGAAGAGAGCTAACTTGCTGA 2718  
Db TGGGTAGAACAGAGAAATTTATAGCAAGTATCCGTTGGAAGAGAGCTAACTTGCTGA 1800  
QY AAAAAAGATTTAAATATTAAGCTATCAATTAAGCAGAAACAGTGAATTTAGTGAAT 2778  
Db AAAAAAGATTTAAATATTAAGCTATCAATTAAGCAGAAACAGTGAATTTAGTGAAT 1860  
QY ATTGAGCTTGGGAAGCAGTATCCGGAAGTCCAGTCCGGAAGGAATTAATAGTTCAA 2838  
Db ATTGAGCTTGGGAAGCAGTATCCGGAAGTCCAGTCCGGAAGGAATTAATAGTTCAA 1920  
QY AGATGCAAAAAATTTAGCTATGCTAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAG 2898  
Db AGATGCAAAAAATTTAGCTATGCTAGAGCTTAAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
QY GGTGGAATGCAATCAATGGAATGGGAAGTATAGTGAAGTATCAGTAAGCAAGAGAT 3018  
Db GGTGGAATGCAATCAATGGAATGGGAAGTATAGTGAAGTATCAGTAAGCAAGAGAT 2100  
QY TCCATTGATGAGAGCAATTTTGAAGCTAATTAAGAGCTAATTAAGAGAGAGAGAGAGAT 3078  
Db TCCATTGATGAGAGCAATTTTGAAGCTAATTAAGAGCTAATTAAGAGAGAGAGAGAGAT 2160  
QY GATACTTCTGTTGGAATGCTGCGGTTCAAGCGGAATCGGAACGAAAAATGCGCGGTC 3138



QY 8545 CAAGGAACAACAAAGCATTTTACGAGATCAAAAGTTAACTTTACAACTGTAAATGGA 8604  
DB 1141 CAAGGAACAACAAAGCATTTTACGAGATCAAAAGTTAACTTTACAACTGTAAATGGA 1200  
QY 8605 GGAACCGATCTCACTTGTGCAAAAGCTTTGGCTTAAATGAAAATTTATGAAATGTA 8664  
DB 1201 GGAACCGATCTCACTTGTGCAAAAGCTTTGGCTTAAATGAAAATTTATGAAATGTA 1260  
QY 8665 AAAGGAACCTGAGAGACCTTAGTCGAGCGGAAACGACGCGCTGAAAAATTTATCAAG 8724  
DB 1261 AAAGGAACCTGAGAGACCTTAGTCGAGCGGAAACGACGCGCTGAAAAATTTATCAAG 1320  
QY 8725 AGTACTACAGAGCATTTGGTTGCAAGAAATTTGGAAATTTGAGATTAATTTAGAAAGAT 8784  
DB 1321 AGTACTACAGAGCATTTGGTTGCAAGAAATTTGGAAATTTGAGATTAATTTAGAAAGAT 1380  
QY 8785 GGAAGAGATTAATACATTTGTAAAGTCAACGAGACGAAACCAAGAGAGCTTTGTCCGA 8844  
DB 1381 GGAAGAGATTAATACATTTGTAAAGTCAACGAGACGAAACCAAGAGAGCTTTGTCCGA 1440  
QY 8845 AAGATGTGTAATTTCTGTGAAAAATACAATTTGAGGGGAAACAAATTCATCTGAAGAT 8904  
DB 1441 AAGATGTGTAATTTCTGTGAAAAATACAATTTGAGGGGAAACAAATTCATCTGAAGAT 1500  
QY 8905 AAAGCCAGAAATTTGTGAAACCGGAAAGTGTAAATGTAGATGCTTTGAAATGAATGATGA 8964  
DB 1501 AAAGCCAGAAATTTGTGAAACCGGAAAGTGTAAATGTAGATGCTTTGAAATGAATGATGA 1560  
QY 8965 GATCTACAAAGAAAAAGTGTGCTATGTGTGAATTTGGTATTTGAAATGTTGATGTAAT 9024  
DB 1561 GATCTACAAAGAAAAAGTGTGCTATGTGTGAATTTGGTATTTGAAATGTTGATGTAAT 1620  
QY 9025 AATGTGATTAAGAAAAATGTAGAGCCAAATCGAAGACATGCTATTGTAGAACTACT 9084  
DB 1621 AATGTGATTAAGAAAAATGTAGAGCCAAATCGAAGACATGCTATTGTAGAACTACT 1680  
QY 9085 GGAACCAAGAAATATCAAGCATTTACAGAGCAAAAGTAAATTTCTTGAAAGAGAGAC 9144  
DB 1681 GGAACCAAGAAATATCAAGCATTTACAGAGCAAAAGTAAATTTCTTGAAAGAGAGAC 1740  
QY 9145 GCTGACAGCTGACGCTGCAATATCGAATGTACATTTCCATGAGATGATTTAAAT 9204  
DB 1741 GCTGACAGCTGACGCTGCAATATCGAATGTACATTTCCATGAGATGATTTAAAT 1800  
QY 9205 TTGGCAAGACAGATGATCTCTCAATTAATACCAAAATTTCAAAAATAATTAATCT 9264  
DB 1801 TTGGCAAGACAGATGATCTCTCAATTAATACCAAAATTTCAAAAATAATTAATCT 1860  
QY 9265 TTAGCATCAAGTAGTGAATCGAATGTGAATGTTCAATGAGGCTGAGCAAGAGTGA 9324  
DB 1861 TTAGCATCAAGTAGTGAATCGAATGTGAATGTTCAATGAGGCTGAGCAAGAGTGA 1920  
QY 9325 GGAAGCAAGAGCAAGTGTGTAAAGATCAAAATTAATGAATTAATTAATTTGATTTA 9384  
DB 1921 GGAAGCAAGAGCAAGTGTGTAAAGATCAAAATTAATGAATTAATTAATTTGATTTA 1980  
QY 9385 GCAAGGAAAAATTAACAGAGGAAACATCAATGTATATGCGGATGATGAATAAATAT 9444  
DB 1981 GCAAGGAAAAATTAACAGAGGAAACATCAATGTATATGCGGATGATGAATAAATAT 2040  
QY 9445 AATATTAAGTAGAACAATTTCTAAGGCTATGCGGATGCGCAAAAGTATGCTGACCTGT 9504  
DB 2041 AATATTAAGTAGAACAATTTCTAAGGCTATGCGGATGCGCAAAAGTATGCTGACCTGT 2100  
QY 9505 TTGGCACTGCGCACTATTGAAAAAATGAAGTAATTTAATTAATGCGATCCGAGAAATTT 9564  
DB 2101 TTGGCACTGCGCACTATTGAAAAAATGAAGTAATTTAATTAATGCGATCCGAGAAATTT 2160  
QY 9565 AAAAAATTAATCTGCAAGATTGGAAGGGAAGCTAATTAATAAATCGTGGTATGAGATCTAAT 9624  
DB 2161 AAAAAATTAATCTGCAAGATTGGAAGGGAAGCTAATTAATAAATCGTGGTATGAGATCTAAT 2220

QY 9625 CAGGTAGACTGTATACGGATTAATATATACATGCACTTTCTGAAAAAAGCATTAACAAAA 9684  
DB 2221 CAGGTAGACTGTATACGGATTAATATATACATGCACTTTCTGAAAAAAGCATTAACAAAA 2280  
QY 9685 TTGACATTAATCAATCAAGAGAGGAGAAAAAAGGAAAAAATGA 9726  
DB 2281 TTGACATTAATCAATCAAGAGAGGAGAAAAAAGGAAAAAATGA 2322

RESULT 8  
US-10-647-057-13  
; Sequence 13, Application US/10647057  
; Publication No. US20040047871A1  
; GENERAL INFORMATION:  
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION  
; APPLICANT: NAGARAJA, T.  
; APPLICANT: STEWART, GEORGE  
; APPLICANT: MARIYAMAN, SANJEEV  
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP  
; FILE REFERENCE: PCT-30962  
; CURRENT APPLICATION NUMBER: US/10/647,057  
; CURRENT FILING DATE: 2003-08-22  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 13  
; LENGTH: 2322  
; TYPE: DNA  
; ORGANISM: Fusobacterium necrophorum  
US-10-647-057-13

Query Match 23.9%; Score 2322; DB 7; Length 2322;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7405 GGAATTAAGCCATTGTGAGAGCTTCTGATATTAATTTGAAAAAATGTAGATATTACACA 7464  
DB 1 GGAATTAAGCCATTGTGAGAGCTTCTGATATTAATTTGAAAAAATGTAGATATTACACA 60  
QY 7465 GAGACCAAGATTAATATCACTTCTACTGCTGTTGGAACTGACAGCTCTGCTCCGCA 7524  
DB 61 GAGACCAAGATTAATATCACTTCTACTGCTGTTGGAACTGACAGCTCTGCTCCGCA 120  
QY 7525 TCAGGACAGTGGCAGTTACAAATATTAAGAAATTTCCGAGTTACTGTGAAATTTCT 7584  
DB 121 TCAGGACAGTGGCAGTTACAAATATTAAGAAATTTCCGAGTTACTGTGAAATTTCT 180  
QY 7585 TTTGTGAAAGAGCTGAAAAAGTAAATGTTAGATCGATATTACAGAAATGTTGCTTTA 7644  
DB 181 TTTGTGAAAGAGCTGAAAAAGTAAATGTTAGATCGATATTACAGAAATGTTGCTTTA 240  
QY 7645 ACAGCATATCAAGTCTCTGTAGAGCATTTGGGAATAGAGCTGCTATGCAATTAAT 7704  
DB 241 ACAGCATATCAAGTCTCTGTAGAGCATTTGGGAATAGAGCTGCTATGCAATTAAT 300  
QY 7705 TCTAATGGAAGATCAAAATATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 7764  
DB 301 TCTAATGGAAGATCAAAATATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 360  
QY 7765 GTTATTGTAAGATTAATTCGAAATTTGAGAGCGGAAAGCAAAAGATTAACCGTAGAGCG 7824  
DB 361 GTTATTGTAAGATTAATTCGAAATTTGAGAGCGGAAAGCAAAAGATTAACCGTAGAGCG 420  
QY 7825 GTAGCTGCCGAGCCATTATCTCAAAAGCAAGATGAATGAATTCAGAGTTGAAATTT 7884  
DB 421 GTAGCTGCCGAGCCATTATCTCAAAAGCAAGATGAATGAATTCAGAGTTGAAATTT 480  
QY 7885 GAGAAGAGTATTTTCAATGAAGAAATAGAGTAATCTAGACCTTCTTAAGAAATTTGAAAGA 7944  
DB 481 GAGAAGAGTATTTTCAATGAAGAAATAGAGTAATCTAGACCTTCTTAAGAAATTTGAAAGA 540  
QY 7945 GAAATCAATGTCAAGTGGAAAAAGAAAAAGAGTGAATGCTGTAATCTCAAGAGCTTCT 8004

Db 541 GAAATCAATGTCGAAAGTGAAGAAAAAGAGAGTCTGAAATCTCCAGAGCTTCT 600  
Qy 8005 GTAGAGACAGTACGAGGGGAGAAATTTTCCGAGCAAAAGATGCCGAGACTTTAT 8064  
Db 601 GTAGAGACAGTACGAGGGGAGAAATTTTCCGAGCAAAAGATGCCGAGACTTTAT 660  
Qy 8065 TTGAAAGTTAGTACAAAATCCGGAAGAAATTTTTCATGCAATATGTAATATGGA 8124  
Db 661 TTGAAAGTTAGTACAAAATCCGGAAGAAATTTTTCATGCAATATGTAATATGGA 720  
Qy 8125 GCACACATTAATTAAGTAAACAGCTTTCTAAGAGTAAAGGTTCTGTATTTGGGA 8184  
Db 721 GCACACATTAATTAAGTAAACAGCTTTCTAAGAGTAAAGGTTCTGTATTTGGGA 780  
Qy 8185 GAGATTGAGTACCAAGGACAGAACTACTGCTGAGGTAATACTATGTAAGTTGAG 8244  
Db 781 GAGATTGAGTACCAAGGACAGAACTACTGCTGAGGTAATACTATGTAAGTTGAG 840  
Qy 8245 GAAAGAAATTTGTCAGAACAATGCAATTTGCAATTTCTAAAGTAAAGTTGAGT 8304  
Db 841 GAAAGAAATTTGTCAGAACAATGCAATTTGCAATTTCTAAAGTAAAGTTGAGT 900  
Qy 8305 GAAAGTAAAGTAACTGCTAAATCTTCTGTAATGTAAGGAAATGAGGAGAAATGCCGA 8364  
Db 901 GAAAGTAAAGTAACTGCTAAATCTTCTGTAATGTAAGGAAATGAGGAGAAATGCCGA 960  
Qy 8365 GAGAGAGTAACTGCTAACTGACAGCAAAAGTAACTGAAATCCGTAAGTTGTTAGGA 8424  
Db 961 GAGAGAGTAACTGCTAACTGACAGCAAAAGTAACTGAAATCCGTAAGTTGTTAGGA 1020  
Qy 8425 CAAAGTTATGAATAATGATTAACAACAAAATATATTTTCAAGTCAATCTCTTGGT 8484  
Db 1021 CAAAGTTATGAATAATGATTAACAACAAAATATATTTTCAAGTCAATCTCTTGGT 1080  
Qy 8485 TTTAATGATACAAAGATGAGCGAATATGAAATCTTTAGCCGTACCGGTGCAATGA 8544  
Db 1081 TTTAATGATACAAAGATGAGCGAATATGAAATCTTTAGCCGTACCGGTGCAATGA 1140  
Qy 8545 CAAAGAAACAAACAAAGCATTTACGAGTCAACAAAGTAACTTCTCAACTGTAAATGA 8604  
Db 1141 CAAAGAAACAAACAAAGCATTTACGAGTCAACAAAGTAACTTCTCAACTGTAAATGA 1200  
Qy 8605 GGAAGCGTATCTCAACTGCTGCAAAAGCTTTGGCTTAAATGAAATATATGAAATGA 8664  
Db 1201 GGAAGCGTATCTCAACTGCTGCAAAAGCTTTGGCTTAAATGAAATATATGAAATGA 1260  
Qy 8665 AAAAGAACTGAGAGAGCTTTAGTCGAGCGGAAACGAGCCGTTGAAATTTACAAAG 8724  
Db 1261 AAAAGAACTGAGAGAGCTTTAGTCGAGCGGAAACGAGCCGTTGAAATTTACAAAG 1320  
Qy 8725 AGTACTACAGAGACATTTGTTGCAAGAAATTTGGGAAATTTGAGATTAATTAAGAAAGAT 8784  
Db 1321 AGTACTACAGAGACATTTGTTGCAAGAAATTTGGGAAATTTGAGATTAATTAAGAAAGAT 1380  
Qy 8785 GCAAGAGATTAATGCAATTTGAGAGTCAACGAGACGGAACCAAGAGAGCTTTGTCGA 8844  
Db 1381 GCAAGAGATTAATGCAATTTGAGAGTCAACGAGACGGAACCAAGAGAGCTTTGTCGA 1440  
Qy 8845 AAGAAATGTAATTTCTGTGAAATAATCAATTTGAGGGGAAACAAATTCATTCATTAAGAT 8904  
Db 1441 AAGAAATGTAATTTCTGTGAAATAATCAATTTGAGGGGAAACAAATTCATTCATTAAGAT 1500  
Qy 8905 AAGCCAGATTTGTTGGAACCGGAAGTAAATGTAATGATCTTTGAATGAATTTGATGA 8964  
Db 1501 AAGCCAGATTTGTTGGAACCGGAAGTAAATGTAATGATCTTTGAATGAATTTGATGA 1560  
Qy 8965 GATCTACAGGAAAGATGCTATGCTATGGAATTTGGAATTTGGAATTTGATGAAT 9024  
Db 1561 GATCTACAGGAAAGATGCTATGCTATGGAATTTGGAATTTGGAATTTGATGAAT 1620  
Qy 9025 AATGTATTAAAGAAATGTAAGACCAAAATCGGAAGACATCTATTGTAAGAACTACT 9084

Db 1621 AATGTATTAAAGAAATGTAAGACCAAAATCGGAAGACATCTATTGTAAGAACTACT 1680  
Qy 9085 GGAAGAAACAAAGATTTCAAGCTTTTACAGAGCAAAAGTAAATTTCTTGGAAAAAGAGAC 9144  
Db 1681 GGAAGAAACAAAGATTTCAAGCTTTTACAGAGCAAAAGTAAATTTCTTGGAAAAAGAGAC 1740  
Qy 9145 GCTGAGCTGAGCTGCAATATGCAATTTGATGACATTTTCAATGAGATGATTAATTAAT 9204  
Db 1741 GCTGAGCTGAGCTGCAATATGCAATTTGATGACATTTTCAATGAGATGATTAATTAAT 1800  
Qy 9205 TTGGCAAGAGTATGCAATCTTCTCAATTAATTAACCAAAATTCAAAAATTAATTAAT 9264  
Db 1801 TTGGCAAGAGTATGCAATCTTCTCAATTAATTAATTAACCAAAATTCAAAAATTAATTAAT 1860  
Qy 9265 TTGAGTATGATGATGATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 9324  
Db 1861 TTGAGTATGATGATGATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCA 1920  
Qy 9325 GAGGCCAAAGGACAGTATGTAAGTAAAGTAAATGAAATGAAATGAAATGAAATGAAAT 9384  
Db 1921 GAGGCCAAAGGACAGTATGTAAGTAAAGTAAATGAAATGAAATGAAATGAAATGAAAT 1980  
Qy 9385 GCAGGAAAAATTTAAACAGAGGAAACATCAATGTAATGCGGATATGATTAATTAAT 9444  
Db 1981 GCAGGAAAAATTTAAACAGAGGAAACATCAATGTAATGCGGATATGATTAATTAAT 2040  
Qy 9445 AATATATGATGATGATTTTCAAGCTTTGCGGATGCGCAAAAGTCAATGCTGCTGCT 9504  
Db 2041 AATATATGATGATGATTTTCAAGCTTTGCGGATGCGCAAAAGTCAATGCTGCTGCT 2100  
Qy 9505 TCGGCACTGCCACTATTTGAAAAAATGAAATTTTAAATGATGATGATGATGATGAT 9564  
Db 2101 TCGGCACTGCCACTATTTGAAAAAATGAAATTTTAAATGATGATGATGATGATGAT 2160  
Qy 9565 AAAAATTAATCTGCAAGATTTGGAAGGAAAGCTAAATTAATTAATTAATTAATTAAT 9624  
Db 2161 AAAAATTAATCTGCAAGATTTGGAAGGAAAGCTAAATTAATTAATTAATTAATTAAT 2220  
Qy 9625 CAGGTAGCTGATATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9684  
Db 2221 CAGGTAGCTGATATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280  
Qy 9685 TTGACATATCAATCAAGAGAGAGGAAAAAGGAAAAATGA 9726  
Db 2281 TTGACATATCAATCAAGAGAGAGGAAAAAGGAAAAATGA 2322

RESULT 9  
US-09-841-786-11  
Sequence 11, Application US/09841786  
Patent No. US2002005483A1  
GENERAL INFORMATION:  
APPLICANT: NAGARAJA, T. G.  
APPLICANT: STEWART, GEORGE C.  
APPLICANT: NARAYANAN, SANJEEV K.  
APPLICANT: CHENGAPPA, M. M.  
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
FILE REFERENCE: 30296  
CURRENT APPLICATION NUMBER: US/09/841,786  
PRIOR FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: 09/558,257  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 11  
LENGTH: 2141  
TYPE: DNA  
ORGANISM: Fusobacterium necrophorum  
US-09-841-786-11

Query Match 22.0%; Score 2141; DB 3; Length 2141;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 3553 GGAGCTTGTGAAGTCTTTCATTAATTAATAAAGGAAACATCGCTGTGTAGAT 3612
Db 1 GGAGCTTGTGAAGTCTTTCATTAATTAATAAAGGAAACATCGCTGTGTAGAT 60
OY 3613 GGAGTAAAGATTAATTTGAAGGAGCAAAATAAAGGTAGAGTGACTTCTGTATCT 3672
Db 61 GGAGTAAAGATTAATTTGAAGGAGCAAAATAAAGGTAGAGTGACTTCTGTATCT 120
OY 3673 ACTTTTGTGGAGCATGGGGCGGATCTGTGCATCTTCACTGAGTAATCATATTGGAAGTGA 3732
Db 121 ACTTTTGTGGAGCATGGGGCGGATCTGTGCATCTTCACTGAGTAATCATATTGGAAGTGA 180
OY 3723 AATAGCAACATCAGTGTGTTAGCTGAGGCGGCTGTAAATTAATTTCAAGTAA 3792
Db 181 AATAGCAACATCAGTGTGTTAGCTGAGGCGGCTGTAAATTAATTTCAAGTAA 240
OY 3793 ACAAGTCTTTGGTTAAAAATAGTATATTGAAAATGCCAATTAATTTAAATAATGCT 3852
Db 241 ACAAGTCTTTGGTTAAAAATAGTATATTGAAAATGCCAATTAATTTAAATAATGCT 300
OY 3853 TTGAGTGAAGAACTCAAGTACAGAGAGAGAGGTTTGAAGCACTTAAAGAAAGTGA 3912
Db 301 TTGAGTGAAGAACTCAAGTACAGAGAGAGAGGTTTGAAGCACTTAAAGAAAGTGA 360
OY 3913 GGACAAGGAAAAAGTTATCTATTGGGAACTTCTGCTTCTATCACTTGTGTGAACATGA 3972
Db 361 GGACAAGGAAAAAGTTATCTATTGGGAACTTCTGCTTCTATCACTTGTGTGAACATGA 420
OY 3973 GTTCTGCAAAATCAGAAAATTAATAAGTAGAGAGAGAACTGAAAGCCAAATAATGAT 4032
Db 421 GTTCTGCAAAATCAGAAAATTAATAAGTAGAGAGAGAACTGAAAGCCAAATAATGAT 480
OY 4033 GTTGAATGCACCTGCTTATCAAGCGGACACCCAGTGAAGAGGCTTAAATTTACAAGCT 4092
Db 481 GTTGAATGCACCTGCTTATCAAGCGGACACCCAGTGAAGAGGCTTAAATTTACAAGCT 540
OY 4093 GGAAGAGTCAAAATGAACTGTAGGGGCTACTGTGACTGTGGCCAAATTTAAACAAGATGA 4152
Db 541 GGAAGAGTCAAAATGAACTGTAGGGGCTACTGTGACTGTGGCCAAATTTAAACAAGATGA 600
OY 4153 AATGCTTCTATTAGTGTGGAGATTAATAAGTTAATCGAAGCGAGCAAAAGCTCTT 4212
Db 601 AATGCTTCTATTAGTGTGGAGATTAATAAGTTAATCGAAGCGAGCAAAAGCTCTT 660
OY 4213 TTAAGAACCACTCAAGTGACTGTGCAAGTGAAGAGGAGCAATTAAGTTCTGAGCG 4272
Db 661 TTAAGAACCACTCAAGTGACTGTGCAAGTGAAGAGGAGCAATTAAGTTCTGAGCG 720
OY 4273 GGAATTAGGAAATTTCAAGGGGCTGTTTCTGTCAATTAAGATTGACAATGACGTGAGCT 4332
Db 721 GGAATTAGGAAATTTCAAGGGGCTGTTTCTGTCAATTAAGATTGACAATGACGTGAGCT 780
OY 4333 AGCGTTGATTAATCTTCCATCGAAGAGCTAAATGAATCAATGATTCGCCAAGATGTC 4392
Db 781 AGCGTTGATTAATCTTCCATCGAAGAGCTAAATGAATCAATGATTCGCCAAGATGTC 840
OY 4393 AAAAGAACTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGAAGATTAATAA 4452
Db 841 AAAAGAACTTCTGATCTAGCAAAAGAAATATCAGGCTTTACTAAATGGAAGATTAATAA 900
OY 4453 TATTTAGAAGATCGTGTATTAAATAGACTGGAATGGTTATTATAGGAAGAAACATGA 4512
Db 901 TATTTAGAAGATCGTGTATTAAATAGACTGGAATGGTTATTATAGGAAGAAACATGA 960
OY 4513 GAAAAAGCAAAAGAAAAAGAGAGCGGTCAATGTAATGCTGCTTTACGTTGCTGGA 4572
Db 961 GAAAAAGCAAAAGAAAAAGAGAGCGGTCAATGTAATGCTGCTTTACGTTGCTGGA 1020
OY 4573 ACGGATTAATCCGCTGAGAGAGTACTATTTGCAAGTCAATCTGTTAAATTAATTTAA 4632
Db 1021 ACGGATTAATCCGCTGAGAGAGTACTATTTGCAAGTCAATCTGTTAAATTAATTTAA 1080
```

```
OY 4633 GCAGAAATTTAGTGAAGCAATTAAGGAGCCGAGAGAGATTAATTCATGCGAAACATGTA 4692
Db 1081 GCAGAAATTTAGTGAAGCAATTAAGGAGCCGAGAGAGATTAATTCATGCGAAACATGTA 1140
OY 4693 AATGTGAGGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGACCTTGCTATCAGC 4752
Db 1141 AATGTGAGGCAAAATCATCTACTGTGTGTGTAATGCGGCTTCTGACCTTGCTATCAGC 1200
OY 4753 AAAAGATCTTTTTCAGGAAATGGGATCTGAGACATGGCAAGACTTATCAATGACAGATT 4812
Db 1201 AAAAGATCTTTTTCAGGAAATGGGATCTGAGACATGGCAAGACTTATCAATGACAGATT 1260
OY 4813 GCAAAAGGTGATTAAGGAAGAAATTTCTGCTGATTCCTTAAATGTAAGCAAAATTAATCC 4872
Db 1261 GCAAAAGGTGATTAAGGAAGAAATTTCTGCTGATTCCTTAAATGTAAGCAAAATTAATCC 1320
OY 4873 ATTCTTGGGGTGAATGTTGCGGGAACCATTCGCGTCTCTTCTACGCGGTAGAGACT 4932
Db 1321 ATTCTTGGGGTGAATGTTGCGGGAACCATTCGCGTCTCTTCTACGCGGTAGAGACT 1380
OY 4933 GCTTTTGGCAATTAATCTCTTCAATTAATAACCTGCTTGTGATTAACGAAACGAAGTA 4992
Db 1381 GCTTTTGGCAATTAATCTCTTCAATTAATAACCTGCTTGTGATTAACGAAACGAAGTA 1440
OY 4993 AATCCTTTAGTGGAAAGATTAACAAAGTCAATGATGACAAAGCTTGAATGATTCATATT 5052
Db 1441 AATCCTTTAGTGGAAAGATTAACAAAGTCAATGATGACAAAGCTTGAATGATTCATATT 1500
OY 5053 ACAAAAGTTTCTGCTGAGAGCGCTGCAAGTATTAAGAGGCTGGAATCGAGAAATGAGTA 5112
Db 1501 ACAAAAGTTTCTGCTGAGAGCGCTGCAAGTATTAAGAGGCTGGAATCGAGAAATGAGTA 1560
OY 5113 TCTGTCAATCGTGTGCTGATGAAACGGAACCTTAAATGATTCGAGTTGGAAGA 5172
Db 1561 TCTGTCAATCGTGTGCTGATGAAACGGAACCTTAAATGATTCGAGTTGGAAGA 1620
OY 5173 GTAAGTCTTTCATATGTAGTGAACAAAGATCAAAAACAAATTAATCAATTCGCCGGAAT 5232
Db 1621 GTAAGTCTTTCATATGTAGTGAACAAAGATCAAAAACAAATTAATCAATTCGCCGGAAT 1680
OY 5233 GCAAAATGAGGAAAAAGCGGCTGAGATTGAGCAACAGTTGCTCATACAAATATTGAAAA 5292
Db 1681 GCAAAATGAGGAAAAAGCGGCTGAGATTGAGCAACAGTTGCTCATACAAATATTGAAAA 1740
OY 5293 CAATCAGTTATAGCTATTTGTAAAAACGTAAATTAACAAGGGAATGATCAAGATGA 5352
Db 1741 CAATCAGTTATAGCTATTTGTAAAAACGTAAATTAACAAGGGAATGATCAAGATGA 1800
OY 5353 AAAAATATCAATGTGACTGCAAAAAGATTAATCTATGACCAATTAATGACAGTGGAGTT 5412
Db 1801 AAAAATATCAATGTGACTGCAAAAAGATTAATCTATGACCAATTAATGACAGTGGAGTT 1860
OY 5413 GGAAGAGCAAAAGAGCGCTGTGCAAGAGCTTCTGCAAGTACTACCTTGAATTAAGCA 5472
Db 1861 GGAAGAGCAAAAGAGCGCTGTGCAAGAGCTTCTGCAAGTACTACCTTGAATTAAGCA 1920
OY 5473 GTTCTTCTCATGTTGATCAAACTGATTAATGACAAAGATTTAGAGAAAGAAATTAATGA 5532
Db 1921 GTTCTTCTCATGTTGATCAAACTGATTAATGACAAAGATTTAGAGAAAGAAATTAATGA 1980
OY 5533 AATTAAGGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATAGAGTCAAGTGCACAAT 5592
Db 1981 AATTAAGGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATAGAGTCAAGTGCACAAT 2040
OY 5593 GCGACAGTGTCTTCCGAGAGCAAGTGAACAAGCTCAGTAGAGAGCTGAGTAGAGATTAA 5652
Db 2041 GCGACAGTGTCTTCCGAGAGCAAGTGAACAAGCTCAGTAGAGAGCTGAGTAGAGATTAA 2100
OY 5653 AAAATTTACAAATTAATCTTCTGCACTATTAATAATATGATAC 5693
Db 2101 AAAATTTACAAATTAATCTTCTGCACTATTAATAATATGATAC 2141
```

```
RESULT 10
US-10-647-057-11
; Sequence 11, Application US/10647057
; Publication No. US20040047871A1
; GENERAL INFORMATION:
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION
; APPLICANT: NAGARAJA, T.
; APPLICANT: STEWART, GEORGE
; APPLICANT: NARAYANAN, SANJEEV
; APPLICANT: CHENGAPPA, M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP
; TITLE OF INVENTION: THEROF
; FILE REFERENCE: PCT-30962
; CURRENT APPLICATION NUMBER: US/10/647,057
; CURRENT FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 2141
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-10-647-057-11

Query Match      22.0%; Score 2141; DB 7; Length 2141;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3553 GGAGGCTTCGGAAGGTTCTTTCAATATATTTAAAGAAACATCTGCTGCTAGAT 3612
DB 1 GGAGCTTCGGAAGGTTCTTTCTTCAATATATTTAAAGAAACATCTGCTGCTAGAT 60
QY 3613 GGAGTAAGATTAATTTGAAGGAGCAATTAAGAGTAGAGTGACTTCTTCTGATTC 3672
DB 61 GGAGTAAGATTAATTTGAAGGAGCAATTAAGAGTAGAGTGACTTCTTCTGATTC 120
QY 3673 ACTTTTGTGGAGCATGGGCGGATCTGCTGCACTTCACTGGAATTTTGAAGTGA 3732
DB 121 ACTTTTGTGGAGCATGGGCGGATCTGCTGCACTTCACTGGAATTTTGAAGTGA 180
QY 3733 AATPAGCATGATGATGCTGTTAGCTGAGCGGCTGCTGTAATTAATTTCAAGTAA 3792
DB 181 AATPAGCATGATGATGCTGTTAGCTGAGCGGCTGCTGTAATTAATTTCAAGTAA 240
QY 3793 ACAAGTCTTTGGTTAAATAATGATATTTGCAATTTGCAATTAATTTAAAGTAA 3852
DB 241 ACAAGTCTTTGGTTAAATAATGATATTTGCAATTTGCAATTAATTTAAAGTAA 300
QY 3853 TTGAGTGAAGAACTCAAGTAGACAGAGCAGGTTTGAAGCAGTTAAAGAAAGTGA 3912
DB 301 TTGAGTGAAGAACTCAAGTAGACAGAGCAGGTTTGAAGCAGTTAAAGAAAGTGA 360
QY 3913 GGAAGAGGAAAGATTAATCTATTTGGAACCTTCTGCTTCAATCACTTAAGAAATGA 3972
DB 361 GGAAGAGGAAAGATTAATCTATTTGGAACCTTCTGCTTCAATCACTTAAGAAATGA 420
QY 3973 GTTTCGCAAAATCAGAAATAATATACAGTAGACAGAGAAATCTGAAAGCCAAATAATGAT 4032
DB 421 GTTTCGCAAAATCAGAAATAATATACAGTAGACAGAGAAATCTGAAAGCCAAATAATGAT 480
QY 4033 GTTGATGTCACTGCTTATCAAGCGGACACCAAGTAGACAGAGCTTTAAATTTTCAAGCT 4092
DB 481 GTTGATGTCACTGCTTATCAAGCGGACACCAAGTAGACAGAGCTTTAAATTTTCAAGCT 540
QY 4093 GGAAGTCAAAATGGAATCTGAGGGGCTACGTGATCTGTTGCCAAATTAACAAAGTA 4152
DB 541 GGAAGTCAAAATGGAATCTGAGGGGCTACGTGATCTGTTGCCAAATTAACAAAGTA 600
QY 4153 AATGCTTCTATTAGTGGTGGAGATATCTAACGTTAATCGACGGAACGCAAAAGCTTT 4212
DB 601 AATGCTTCTATTAGTGGTGGAGATATCTAACGTTAATCGACGGAACGCAAAAGCTTT 660
QY 4213 TTAGCAACCACTCAAGTACTGCTGCAGTGACAGCGGAGGAGCAATTAGTTCTGAGAGC 4272
```

```
DB 661 TTAGCAACCACTCAAGTACTGCTGCAGTGACAGCGGAGGAGCAATTAGTTCTGAGAGC 720
QY 4273 GGATTAGGAATTTATCAAGGGGCTGTTCTGCTCAATTAAGATTGACATGACGTGAAGCT 4332
DB 721 GGATTAGGAATTTATCAAGGGGCTGTTCTGCTCAATTAAGATTGACATGACGTGAAGCT 780
QY 4333 AGCGTTGATAATCTTCATCGAAGGAGCTAATGAATCAATGATTCAGCAAGATGTC 4392
DB 781 AGCGTTGATAATCTTCATCGAAGGAGCTAATGAATCAATGATTCAGCAAGATGTC 840
QY 4393 AAAGGAATTCGATCTAGCAAAAGATATATGAGCTTTACTAATATGAAAAATAAAA 4452
DB 841 AAAGGAATTCGATCTAGCAAAAGATATATGAGCTTTACTAATATGAAAAATAAAA 900
QY 4453 TATTAGGAAGATCGTGATTATATACGCTGAAATGTTATATACGAGAACACTA 4512
DB 901 TATTAGGAAGATCGTGATTATATACGCTGAAATGTTATATACGAGAACACTA 960
QY 4513 GAAAAAGCAAGAAAAAGAGCGGCTCAATGCTTAAATGCTTAAATGCTTAAATGCTGGA 4572
DB 961 GAAAAAGCAAGAAAAAGAGCGGCTCAATGCTTAAATGCTTAAATGCTTAAATGCTGGA 1020
QY 4573 ACGGATAATCCGCTGAGAGTAGCTATTGCAATTAATGCTTAAATGCTTAAATGCTTAA 4632
DB 1021 ACGGATAATCCGCTGAGAGTAGCTATTGCAATTAATGCTTAAATGCTTAAATGCTTAA 1080
QY 4633 GCAGATTAAGTGAAGCAATTAAGAACCCGAGAGATTAATTAATCATGCGAATGTA 4692
DB 1081 GCAGATTAAGTGAAGCAATTAAGAACCCGAGAGATTAATTAATCATGCGAATGTA 1140
QY 4693 AATGAGAGGCAAAATCTATCTGTTGTTGGAATGCGGCTTCTGAGCTTCTATCAGC 4752
DB 1141 AATGAGAGGCAAAATCTATCTGTTGTTGGAATGCGGCTTCTGAGCTTCTATCAGC 1200
QY 4753 AAAGATGCTTTTTCAGAAATGGAGATCTGAGCATGCAAGCTTATCAATGACAGAT 4812
DB 1201 AAAGATGCTTTTTCAGAAATGGAGATCTGAGCATGCAAGCTTATCAATGACAGAT 1260
QY 4813 GCAGAGTGAATTAAGAGAAATTTCTGCTGATCTTAAATGTAAGCAATTAATTC 4872
DB 1261 GCAGAGTGAATTAAGAGAAATTTCTGCTGATCTTAAATGTAAGCAATTAATTC 1320
QY 4873 ATTCTTGGGATTAATGTTGCGGGAACATTTGCGGCTTCTTCTTCAAGCGGCTAGAGCT 4932
DB 1321 ATTCTTGGGATTAATGTTGCGGGAACATTTGCGGCTTCTTCTTCAAGCGGCTAGAGCT 1380
QY 4933 GCTTTTGGGAATTAATGCTTCAATATAAATCTGCTTGAATTAACAGAAAGTA 4992
DB 1381 GCTTTTGGGAATTAATGCTTCAATATAAATCTGCTTGAATTAACAGAAAGTA 1440
QY 4993 AATCCTTTTATGAGAAAGATACAAAGTCAATGTAACAGCTTTGAATGATTTCTCATATT 5052
DB 1441 AATCCTTTTATGAGAAAGATACAAAGTCAATGTAACAGCTTTGAATGATTTCTCATATT 1500
QY 5053 ACAAAAGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGAATCGAGAAATGTA 5112
DB 1501 ACAAAAGTTTCTGCTGAGGCGCTGCAAGTATTAAGCAGGCTGGAATCGAGAAATGTA 1560
QY 5113 TCTGTCAATCGGTTCTGATGAAACGGAAGCTTATGATGATTTCTGATTTGAAGTA 5172
DB 1561 TCTGTCAATCGGTTCTGATGAAACGGAAGCTTATGATGATTTCTGATTTGAAGTA 1620
QY 5173 GTAAGTCTTTTCAATGATGATCAAAAGATCAAAAGATCAATTAATCAATTTCCGGAAT 5232
DB 1621 GTAAGTCTTTTCAATGATGATCAAAAGATCAAAAGATCAATTAATCAATTTCCGGAAT 1680
QY 5233 GCAATGAGAGAAAGCGGCTGAGTGGAGCAAGTTGCTCATACAAATTTTGAAGTA 5292
DB 1681 GCAATGAGAGAAAGCGGCTGAGTGGAGCAAGTTGCTCATACAAATTTTGAAGTA 1740
QY 5293 CAATGATTAATGTTTAAAGAAACAGTAATAATTAACAAGCGAATGATCAAGTAGA 5352
```

Db 1741 CAATAGTTATAGCTATTGTGTAACAAAGTAATAATTACACGGCGAATGATCAGATAGCA 1800  
Qy 5353 AAAAATATCAATGTGACTGCAAAAGATTATATCTATGACCAATACTATAGACGTCCGAGTT 5412  
Db 1801 AAAAATATCAATGTGACTGCAAAAGATTATATCTATGACCAATACTATAGACGTCCGAGTT 1860  
Qy 5413 GGAGAGCAAAAGAGGCTCTGTGCAAGAGCTTCTGCAAGTACTTACCTTGAATTAAGCA 5472  
Db 1861 GGAGAGCAAAAGAGGCTCTGTGCAAGAGCTTCTGCAAGTACTTACCTTGAATTAAGCA 1920  
Qy 5473 GTTCTCTGCAAGTGTGATCAAACTGATATGACAAAGATTGAGAGAGAAATAATAGCA 5532  
Db 1921 GTTCTCTGCAAGTGTGATCAAACTGATATGACAAAGATTGAGAGAGAAATAATAGCA 1980  
Qy 5533 AATAAGAAAAGCAAAATGTATATGTTCTAGCTGAAAATACGAGTCAAGTGGTCACAAT 5592  
Db 1981 AATAAGAAAAGCAAAATGTATATGTTCTAGCTGAAAATACGAGTCAAGTGGTCACAAT 2040  
Qy 5593 GCGACAGTGTCTTCCGAGCAAGTGACAAAGCTGCAAGTGAAGAGCTGAGTAGCAGTTAAT 5652  
Db 2041 GCGACAGTGTCTTCCGAGCAAGTGACAAAGCTGCAAGTGAAGAGCTGAGTAGCAGTTAAT 2100  
Qy 5653 AAAATTACACAAAATTAATCTTCTGACATATAAAAAATAGTAC 5693  
Db 2101 AAAATTACACAAAATTAATCTTCTGACATATAAAAAATAGTAC 2141

RESULT 11  
US-09-841-786-12  
Sequence 12, Application US/09841786  
Patent No. US2002005483A1  
GENERAL INFORMATION:  
APPLICANT: NAGARAJA, T. G.  
APPLICANT: STEWART, GEORGE C.  
APPLICANT: NARAYANAN, SANJEEV K.  
APPLICANT: CHENGAPPA, M. M.  
TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN  
TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF  
FILE REFERENCE: 30296  
CURRENT APPLICATION NUMBER: US/09/841,786  
CURRENT FILING DATE: 2001-04-24  
PRIOR APPLICATION NUMBER: 09/558,257  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 12  
LENGTH: 1887  
TYPE: DNA  
ORGANISM: Fusobacterium necrophorum  
US-09-841-786-12

Query Match 19.4%; Score 1887; DB 3; Length 1887;  
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;  
Matches 1887; Conservative 0;

Qy 5624 CTGAGTAGAGAGCTGAGTAGAGAGTTAATTAATTAACCAAAATATCTTCTGACATATTA 5683  
Db 1 CTGAGTAGAGAGCTGAGTAGAGAGTTAATTAATTAACCAAAATATCTTCTGACATATTA 60  
Qy 5684 AAAATAGTACTCAAAATGTACGAATGCTTTGGTAAAAAGCAATCTCATCTATTA 5743  
Db 61 AAAATAGTACTCAAAATGTACGAATGCTTTGGTAAAAAGCAATCTCATCTATTA 120  
Qy 5744 AAAAATTGGAATTTGAGAGCTGAGTTGAGCTGAGAGAGCTGAGTAGCAGGTTCTGAG 5803  
Db 121 AAAAATTGGAATTTGAGAGCTGAGTTGAGCTGAGAGAGCTGAGTAGCAGGTTCTGAG 180  
Qy 5804 CAGTGAATTAAGATTGTAATTAATAGCATAGCAGAAATTAATCATCAAAATCAGTCGCA 5863  
Db 181 CAGTGAATTAAGATTGTAATTAATAGCATAGCAGAAATTAATCATCAAAATCAGTCGCA 240  
Qy 5864 AGGAAATGTGAGAGTTATTACAGAGTCTGATGCGGTAAATTCATTAATTAGCAGAAACAG 5923

Db 241 AGGAAATGTGAGAGTTATTACAGAGTCTGATGCGGTAAATTCATTAATTAGCAGAAACAG 300  
Qy 5924 TGTCTGAGTGGCCCGTGCAGCAATATAGAGGCTCAACAGATGTGAATGAATTAACAGAT 5983  
Db 301 TGTCTGAGTGGCCCGTGCAGCAATATAGAGGCTCAACAGATGTGAATGAATTAACAGAT 360  
Qy 5984 CTACAAAAGCATATGTAAAGATTCTACAGTGAATTTGCTAAAGAAAGAAACAGATGATTA 6043  
Db 361 CTACAAAAGCATATGTAAAGATTCTACAGTGAATTTGCTAAAGAAAGAAACAGATGATTA 420  
Qy 6044 TTACTACTCAGGGCAGTATATTAAGTGTAGATTAAGATTAAGATTTCAAAAATCTTAATTA 6103  
Db 421 TTACTACTCAGGGCAGTATATTAAGTGTAGATTAAGATTTCAAAAATCTTAATTA 480  
Qy 6104 ACGAAGACTTATCACAAAAAAGAAAAATAGTAATTAATAAAAAAGATTGTACCAATAGTT 6163  
Db 481 ACGAAGACTTATCACAAAAAAGAAAAATAGTAATTAATAAAAAAGATTGTACCAATAGTT 540  
Qy 6164 CAGTACTCATATCTTTAAATCTTTATTTGGCAAAATGCCGCTGTTCAGAGCAAGCCGAG 6223  
Db 541 CAGTACTCATATCTTTAAATCTTTATTTGGCAAAATGCCGCTGTTCAGAGCAAGCCGAG 600  
Qy 6224 TGGCAGAAAGCTGTTAATATCAACAAGTTTATGAGAAACAGAAAGCTTTGTAGAAAAT 6283  
Db 601 TGGCAGAAAGCTGTTAATATCAACAAGTTTATGAGAAACAGAAAGCTTTGTAGAAAAT 660  
Qy 6284 CTATATTAAATGCAAAAACATTAATCTGTAAATACAGAGATTAACAGAAATCAATCCGAG 6343  
Db 661 CTATATTAAATGCAAAAACATTAATCTGTAAATACAGAGATTAACAGAAATCAATCCGAG 720  
Qy 6344 TAGTAGGTTCTGTTGGTGTGTGAAAATGTAGAGTAGAGCTTTCTGTATACAAATTA 6403  
Db 721 TAGTAGGTTCTGTTGGTGTGTGAAAATGTAGAGTAGAGCTTTCTGTATACAAATTA 780  
Qy 6404 TTATTAATAAGAAATATCAACAAGCAAGAGTTGAAAAAACTACAAATGTCTGATGAAGTTCCG 6463  
Db 781 TTATTAATAAGAAATATCAACAAGCAAGAGTTGAAAAAACTACAAATGTCTGATGAAGTTCCG 840  
Qy 6464 GAGAAGAAAGCTGAATTAACAGCAAGATTTCTAAGCAAGAAATTTCTTTGAGAGTCGAG 6523  
Db 841 GAGAAGAAAGCTGAATTAACAGCAAGATTTCTAAGCAAGAAATTTCTTTGAGAGTCGAG 900  
Qy 6524 TCGAGCAGCCGGGATAGAGCCGAGAGTGAGAGAAACGGTTCCGTAAATCAATTTGAG 6583  
Db 901 TCGAGCAGCCGGGATAGAGCCGAGAGTGAGAGAAACGGTTCCGTAAATCAATTTGAG 960  
Qy 6584 GAAAGACGGAAGTAGATGTGAAAGCAAGATTTTGTGTAATAAAAAAGCTGAGATTACAG 6643  
Db 961 GAAAGACGGAAGTAGATGTGAAAGCAAGATTTTGTGTAATAAAAAAGCTGAGATTACAG 1020  
Qy 6644 CAAAACGTTATAGTTCTGTTGCAATTTGGAATTTGCCAGCTGCGAGTGGCTCAAAAGAG 6703  
Db 1021 CAAAACGTTATAGTTCTGTTGCAATTTGGAATTTGCCAGCTGCGAGTGGCTCAAAAGAG 1080  
Qy 6704 CTGGAATTTGAGCAGCAGTGGCAGTTACCAAGATGAATCAAAACAGAGCAAGAGTGA 6763  
Db 1081 CTGGAATTTGAGCAGCAGTGGCAGTTACCAAGATGAATCAAAACAGAGCAAGAGTGA 1140  
Qy 6764 AAAATTCGTAATTAATGACTGGAACCAAGTTAGTGAATAGCAGAAAAATGAGTAAT 6823  
Db 1141 AAAATTCGTAATTAATGACTGGAACCAAGTTAGTGAATAGCAGAAAAATGAGTAAT 1200  
Qy 6824 CAGGTACTGGAATCGGTTACGCCGAGAGCTGGAATTTCTTGACGCCGAGATCTGGAAGTGG 6883  
Db 1201 CAGGTACTGGAATCGGTTACGCCGAGAGCTGGAATTTCTTGACGCCGAGATCTGGAAGTGG 1260  
Qy 6884 TTTCTGTCATTAATATTTGCAAAATAGGTGAACAGATATGATCATATGATCTTTACACT 6943  
Db 1261 TTTCTGTCATTAATATTTGCAAAATAGGTGAACAGATATGATCATATGATCTTTACACT 1320  
Qy 6944 CTTTACTGATGTAAATGTAAAGCTCTTAATTAATTTGCAATTTCTTGACAGCCGGTG 7003  
Db 1321 CTTTACTGATGTAAATGTAAAGCTCTTAATTAATTTGCAATTTCTTGACAGCCGGTG 1380



```
|||||
Db 1201 CAGGACTGGAAATCGGTTGACGCCGAGCTGGAATCTTTCACGCCGAGATCTGGAGTGG 1260
Qy 6884 TTTCTGCTAATTAATTTGCAATTAAGGTGAAACAGATTCATGATCTTAACT 6943
Db 1261 TTTCTGCTAATTAATTTGCAATTAAGGTGAAACAGATTCATGATCTTAACT 1320
Qy 6944 CTTCTACTGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 7003
Db 1321 CTTCTACTGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1380
Qy 7004 GAGGAGCCGAGGTCTTGACAGAGTTACCGAGTGGTTCTGTTAACTATTAATGTT 7063
Db 1381 GAGGAGCCGAGGTCTTGACAGAGTTACCGAGTGGTTCTGTTAACTATTAATGTT 1440
Qy 7064 CTGTGATAGCTCGAGTTCAATTAATCTGATTTGACTTCGCTAGAGAAAAGTAATG 7123
Db 1441 CTGTGATAGCTCGAGTTCAATTAATCTGATTTGACTTCGCTAGAGAAAAGTAATG 1500
Qy 7124 TAACGCAAAAGAGAAAATATTAAGCAAAACAGACAAATGCAAGATCGAGAG 7183
Db 1501 TAACGCAAAAGAGAAAATATTAAGCAAAACAGACAAATGCAAGATCGAGAG 1560
Qy 7184 CAGCAATCGAGCAATGTTCTTGTAAATTAATTTTGGAAACAGCTGTAGAAATAGAAAA 7243
Db 1561 CAGCAATCGAGCAATGTTCTTGTAAATTAATTTTGGAAACAGCTGTAGAAATAGAAAA 1620
Qy 7244 ATTCTGAAGAAAAGAGACAGAAAGTTTAAACCTTTAGACGAATGTAACAAAGAACAG 7303
Db 1621 ATTCTGAAGAAAAGAGACAGAAAGTTTAAACCTTTAGACGAATGTAACAAAGAACAG 1680
Qy 7304 ATTAATAAGTAATGATGCTACGAAAATCTTCAATCAGCAGATTTTCTACAGAG 7363
Db 1681 ATTAATAAGTAATGATGCTACGAAAATCTTCAATCAGCAGATTTTCTACAGAG 1740
Qy 7364 ATACTTCTGTAAGAGCGATAGAGAGATTAAGAGAGAGAAATTAAGCCATTGTA 7423
Db 1741 ATACTTCTGTAAGAGCGATAGAGAGATTAAGAGAGAGAAATTAAGCCATTGTA 1800
Qy 7424 AACCTTCTGATATTAATTTGAAAAAATGTGATATTAACACAGAGACAAAGTAATATCA 7483
Db 1801 AACCTTCTGATATTAATTTGAAAAAATGTGATATTAACACAGAGACAAAGTAATATCA 1860
Qy 7484 CTTCTACTGATGTTGGAACTGCAG 7510
Db 1861 CTTCTACTGATGTTGGAACTGCAG 1887

RESULT 13
US-09-841-786-9
; Sequence 9, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO: 9
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-9
```

Query Match 11.6%; Score 1130; DB 3; Length 1130;

```
Best Local Similarity 100.0%; Pred. No. 8,9e-187;
Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGCGGCTACCAAAATTAACGTTCAAGAGACAAAGAAAGAGATTCGATTTCTAAAAA 60
Db 1 ATGAGCGGCTACCAAAATTAACGTTCAAGAGACAAAGAAAGAGATTCGATTTCTAAAAA 60
Qy 61 GTTTTAATGATTTTGGAGATGTTGATTAACTATGACGTGAGAGGCTATGATCAATC 120
Db 61 GTTTTAATGATTTTGGAGATGTTGATTAACTATGACGTGAGAGGCTATGATCAATC 120
Qy 121 ACCGCACTGAGAAATTTTGGAAACAAAATAGAAAAAAGATTAATGTTATGACATTA 180
Db 121 ACCGCACTGAGAAATTTTGGAAACAAAATAGAAAAAAGATTAATGTTATGACATTA 180
Qy 181 ACCAAACAAGATTCAGAGGAGAAAGCTTTTAAAGTTTAAATGATTTGCTTTAAACAG 240
Db 181 ACCAAACAAGATTCAGAGGAGAAAGCTTTTAAAGTTTAAATGATTTGCTTTAAACAG 240
Qy 241 AATAATATAGCAAAATCTATATTTTGGGAAAAGATAGTACGGGGGTAAATATCTTTT 300
Db 241 AATAATATAGCAAAATCTATATTTTGGGAAAAGATAGTACGGGGGTAAATATCTTTT 300
Qy 301 AACTTGTCAATGGAATAATTTGAAGTAGATGGAATTATCAACGAAATTCAGAAAAATAA 360
Db 301 AACTTGTCAATGGAATAATTTGAAGTAGATGGAATTATCAACGAAATTCAGAAAAATAA 360
Qy 361 ATTGAGAGAAATTTATATTTTCTTAAGCTCGAGAAAGGATGCGAGTAGAAAAATAGAG 420
Db 361 ATTGAGAGAAATTTATATTTTCTTAAGCTCGAGAAAGGATGCGAGTAGAAAAATAGAG 420
Qy 421 ATCAATGCTGTTCTTTTCAATTTCTATTAATTCACAAAACAAGATGATTTTAAAGGCT 480
Db 421 ATCAATGCTGTTCTTTTCAATTTCTATTAATTCACAAAACAAGATGATTTTAAAGGCT 480
Qy 481 GAAGAGCCAAACATGTGTAAGTTTAAATGAAATCAATTCAGATAGATGAAAAAGTAAA 540
Db 481 GAAGAGCCAAACATGTGTAAGTTTAAATGAAATCAATTCAGATAGATGAAAAAGTAAA 540
Qy 541 ATTCATTTGAATCCGAATGGAAGCATTCGGTAGAAGAAAATCAATGCTGTTGAAAGC 600
Db 541 ATTCATTTGAATCCGAATGGAAGCATTCGGTAGAAGAAAATCAATGCTGTTGAAAGC 600
Qy 601 ATCGGTTTATATGCGCGGAGATTAATGATTAAGATTAAGATTAAGATTAAGATTAAG 660
Db 601 ATCGGTTTATATGCGCGGAGATTAATGATTAAGATTAAGATTAAGATTAAGATTAAG 660
Qy 661 ACAGATTTTAAAAATTTAGTCAATTAATGATGGAATTAATTTCTGCTGACCGAGAT 720
Db 661 ACAGATTTTAAAAATTTAGTCAATTAATGATGGAATTAATTTCTGCTGACCGAGAT 720
Qy 721 TTTAAAGCTACCAAGCAAAATCTGAGATTTATTTCTTTAGCTCACATGATTTCTCT 780
Db 721 TTTAAAGCTACCAAGCAAAATCTGAGATTTATTTCTTTAGCTCACATGATTTCTCT 780
Qy 781 CAAAAGCTATGAGGAAAAATTTCAACTGTTGAAAAGAGAAATAGAAATATGTAAGAA 840
Db 781 CAAAAGCTATGAGGAAAAATTTCAACTGTTGAAAAGAGAAATAGAAATATGTAAGAA 840
Qy 841 AATAACCAAGCAAAATTTGATCTGATGCTGATTTGGAAGAGATGGAATATTAATAAT 900
Db 841 AATAACCAAGCAAAATTTGATCTGATGCTGATTTGGAAGAGATGGAATATTAATAAT 900
Qy 901 AGTGGAAAGCTACCAATGGAAGATTTTAAAGAAAGAGGAGAAAAAGAACTTATTAAC 960
Db 901 AGTGGAAAGCTACCAATGGAAGATTTTAAAGAAAGAGGAGAAAAAGAACTTATTAAC 960
Qy 961 ACTCCTTAATGTTATGATGATGGAAGCTCGTAAGATTAATTAAGAAAGAAAGTCA 1020
Db 961 ACTCCTTAATGTTATGATGATGGAAGCTCGTAAGATTAATTAAGAAAGAAAGTCA 1020
Qy 1021 GAAAAGATGTTGACATTTACGCTGAAGCAAAAGATTTCTATGATGCAATTTAGTTACT 1080
Db 1021 GAAAAGATGTTGACATTTACGCTGAAGCAAAAGATTTCTATGATGCAATTTAGTTACT 1080
```

Db 1021 GGAAGAATGTTGACATTACAGCTGAGCAAGAAATTTCTAATGATGCAACTTACTACT 1080  
Qy 1081 AAGCTTGCAAAAGCACTTTTACGTTTGTGTTACAGGTTCTATTCTCTAT 1130  
Db 1081 AAGCTTGCAAAAGCACTTTTACGTTTGTGTTACAGGTTCTATTCTCTAT 1130  
RESULT 14  
US-10-647-057-9  
; Sequence 9, Application US/10647057  
; Publication No. US20040047871A1  
; GENERAL INFORMATION:  
; APPLICANT: KANSAS STATE UNIVERSITY RESEARCH FOUNDATION  
; APPLICANT: NAGARAJA, T.  
; APPLICANT: STEWART, GEORGE  
; APPLICANT: NARAYANAN, SANJEEV  
; APPLICANT: CHENGAPPA, M.  
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN VACCINE AND PREP  
; TITLE OF INVENTION: THEROPF  
; FILE REFERENCE: PCT-30962  
; CURRENT APPLICATION NUMBER: US/10/647,057  
; CURRENT FILING DATE: 2003-08-22  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 1130  
; TYPE: DNA  
; ORGANISM: Fusobacterium necrophorum  
US-10-647-057-9  
Query Match 11.6%; Score 1130; DB 7; Length 1130;  
Best Local Similarity 100.0%; Pred. No. 8.9e-187; Indels 0; Gaps 0;  
Matches 1130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 1 ATGAGCGGATCAAAATATACGTTGAGAGACAGAGAAAGGATATACAGATTCTAATAAAA 60  
Db 1 1 ATGAGCGGATCAAAATATACGTTGAGAGACAGAGAAAGGATATACAGATTCTAATAAAA 60  
Qy 61 GTTTTATGATTTGGGATTTGTTGATTAACATATGACGGTGAAGGCTTAATGATCAATC 120  
Db 61 GTTTTATGATTTGGGATTTGTTGATTAACATATGACGGTGAAGGCTTAATGATCAATC 120  
Qy 121 ACCGCGATCGAATTTTGGAAACAAATAAGAAAAAGATTAATGATTAATGATTAAT 180  
Db 121 ACCGCGATCGAATTTTGGAAACAAATAAGAAAAAGATTAATGATTAATGATTAAT 180  
Qy 181 ACAAACAAGATCAAGGGGAGAAAGCTTTTACAGTTTAATGATTAATGATTAATGATTA 240  
Db 181 ACAAACAAGATCAAGGGGAGAAAGCTTTTACAGTTTAATGATTAATGATTAATGATTA 240  
Qy 241 AATAATATAGCAATCTATATTTTGGGAAAAAGATAGTACGGGGTAAATATCTTTT 300  
Db 241 AATAATATAGCAATCTATATTTTGGGAAAAAGATAGTACGGGGTAAATATCTTTT 300  
Qy 301 AACTTTGTCATGAAAAATGAGTAGATGCGATTATCAACGGAATTCGAGAAATATA 360  
Db 301 AACTTTGTCATGAAAAATGAGTAGATGCGATTATCAACGGAATTCGAGAAATATA 360  
Qy 361 ATTGAGGAAATTTTATTTCTTAAGCTGGAAAGGATGCGATAGAGAAAAATGAGATT 420  
Db 361 ATTGAGGAAATTTTATTTCTTAAGCTGGAAAGGATGCGATAGAGAAAAATGAGATT 420  
Qy 421 ATCAATGCTGTTCTTTCTATCTATATTTCCAAAAAGATGATTTTAAAGAGCTTTG 480  
Db 421 ATCAATGCTGTTCTTTCTATCTATATTTCCAAAAAGATGATTTTAAAGAGCTTTG 480  
Qy 481 GAAGAAAGCCAAACATGTAAGTTTAAATGAAATATTCAGTAGATGAAAAAGTAAAA 540  
Db 481 GAAGAAAGCCAAACATGTAAGTTTAAATGAAATATTCAGTAGATGAAAAAGTAAAA 540  
Qy 541 ATTCCATTGAATCCGAATGAGCATTAACGTTAGAGGAAAGAAATCAATGCTTTAGAGG 600  
Db 541 ATTCCATTGAATCCGAATGAGCATTAACGTTAGAGGAAAGAAATCAATGCTTTAGAGG 600

Qy 601 ATCGGTTATATGCGGCGGATATTAGATTGAAAGATACGCAATCTAATAAGACAGAAAT 660  
Db 601 ATCGGTTATATGCGGCGGATATTAGATTGAAAGATACGCAATCTAATAAGACAGAAAT 660  
Qy 661 ACAGATTTTAAATTTTATGCAATATTTAGTATGATGCAATTAATTTCTGCTGACCGAGAT 720  
Db 661 ACAGATTTTAAATTTTATGCAATATTTAGTATGATGCAATTAATTTCTGCTGACCGAGAT 720  
Qy 721 TTTAAAGCTTACCAAGACAAATCTGAGATATTTATTTCTTACGCTCAGATGATTTCTCT 780  
Db 721 TTTAAAGCTTACCAAGACAAATCTGAGATATTTATTTCTTACGCTCAGATGATTTCTCT 780  
Qy 781 CAAAAGCTATGCGGAAAAATTTCACTGTGGAAGAAATGAAATATGTAAGAA 840  
Db 781 CAAAAGCTATGCGGAAAAATTTCACTGTGGAAGAAATGAAATATGTAAGAA 840  
Qy 841 AATACCAAGCAATATTTGATCTGATCTGATTTGAAAGCAGATGAAATTAATAAAT 900  
Db 841 AATACCAAGCAATATTTGATCTGATCTGATTTGAAAGCAGATGAAATTAATAAAT 900  
Qy 901 AGTGGAAGCTACCAATGGAATTTTAAAGAAAGGGAAGAAAGAACTTATTAAC 960  
Db 901 AGTGGAAGCTACCAATGGAATTTTAAAGAAAGGGAAGAAAGAACTTATTAAC 960  
Qy 961 ACTCTTTTATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 961 ACTCTTTTATGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Qy 1021 GGAAGAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTACTACT 1080  
Db 1021 GGAAGAATGTTGACATTACAGCTGAGCAAGAAATTTCTATGATGCAACTTACTACT 1080  
Qy 1081 AAGCTTGCAAAAGCACTTTTACGTTTGTGTTACAGGTTCTATTCTCTAT 1130  
Db 1081 AAGCTTGCAAAAGCACTTTTACGTTTGTGTTACAGGTTCTATTCTCTAT 1130  
RESULT 15  
US-09-742-096-2  
; Sequence 2, Application US/09742096  
; Patent No. US20020155441A1  
; GENERAL INFORMATION:  
; APPLICANT: DRUIHE, PIERRE  
; APPLICANT: DAUBERLIES, PIERRE  
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES  
; FILE REFERENCE: 200773USODIV  
; CURRENT APPLICATION NUMBER: US/09/742,096  
; CURRENT FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 08/973,642  
; PRIOR FILING DATE: 1998-02-06  
; PRIOR APPLICATION NUMBER: PCT/FR96/00894  
; PRIOR FILING DATE: 1996-06-12  
; PRIOR APPLICATION NUMBER: FR 95/07007  
; PRIOR FILING DATE: 1995-06-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 5361  
; TYPE: DNA  
; ORGANISM: P. falciparum  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(5361)  
; OTHER INFORMATION:  
US-09-742-096-2  
Query Match 1.2%; Score 116.2; DB 3; Length 5361;  
Best Local Similarity 41.7%; Pred. No. 9e-10; Indels 67; Gaps 14;  
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;  
Qy 5126 GTTCGATGAACGGAAGCTTAGTTAGTATTCGAGTTGAAGAGTAAGTCTTTCA 5185  
Db 5126 GTTCGATGAACGGAAGCTTAGTTAGTATTCGAGTTGAAGAGTAAGTCTTTCA 5185

Db 773 GTGTGAAGAAAGTATAGCTTCAAGTGTGATGAAGAAATATAGATTCAAGTATTTGAAGAA 832  
Qy 5186 ATGTAGATGCAAAAGATCAAAACATTAATACAAA---TTGCCGGAATGCAATGAG 5242  
Db 833 ATGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAAAGTGTGTGAAGAAATGTGGCTCCAA 892  
Qy 5243 GAAAACCGCTGAGATTGAGACACAGTTGCTCATACAAATATTGAAAAACATCAGTTA 5302  
Db 893 GTGTGAAGAAAGTGTAGAAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTGGAAGAA 952  
Qy 5303 TAGCATTTGTAATAAAACAGTAAATTTACACGCGCAATGATCAAGATGAAAAATATCA 5362  
Db 953 GTTAGCTGAAAAGTGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1012  
Qy 5363 ATGTAGCTGCAAAAGATTATATCTATGACCAATACTATAGCAGTCCGAGTTGAGAGACAA 5422  
Db 1013 ATGTGAAGAAATCGTACTCCAACTGTGGAAGAAATGTGTAGCTCCAACTGTGGAAGAA 1072  
Qy 5423 AAGAGCTCTGTGCAAGAGCTTCTGCAAGTACTACTGTGAATGACAGTTTCTTCTC 5482  
Db 1073 TTGTAGCTCCAAAGTGTGTAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAGAA 1132  
Qy 5483 ATGTGTATCAAAAGTATTTGACAAAGATTTAGAGAGAAAGAAATATAGAAATTAAGAAA 5542  
Db 1133 ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA 1192  
Qy 5543 AAGCAATGTTATATGTTCTAGCTGAAAATAGCAGTCAAGTGTGTCACAAATGCGACAGTGC 5602  
Db 1193 GTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1252  
Qy 5603 TTTCCGAGCAAGTGTGACAGAGCTGTGAGAGCTGTGAGAGCTGATTAATTAATTTACAC 5662  
Db 1253 ATGTGAAGAAATCGTACTCCAACTGTGGAAGAAATCGTAGCTCCAACTGTGGAAGAA 1312  
Qy 5663 AAAATCTCTGCAATATAAAAAATAGTACTCAAAATGTGACGAAATCTTGTGTAATAA 5722  
Db 1313 TTGTAGCTCCAAAGTGTGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTAGAGAA 1372  
Qy 5723 GC---AAATCTCATTCATCTATTAACAATTTGCAATTTGAGCTGAGATTGAGAGCTGAG 5779  
Db 1373 ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA 1432  
Qy 5780 GAGCTGAGTACAGAGTTCTGTAGCAGTGAATAGATTGTAAATTAATTCGATAGCAGAT 5839  
Db 1433 GTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1492  
Qy 5840 TAAATCATGCAAAAAATCACTGCAAGGGAATGTGCGAGTTATTAACAAGCTGATGCGG 5899  
Db 1493 ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA 1552  
Qy 5900 TAAATGCTAATTTATGACAGAAACAGTGTCTGAGAGTGGCCGTCGACGCAATAGAGCTCAA 5959  
Db 1553 TCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATTTGAGCTCAA 1612  
Qy 5960 CCAAGTGAATGAATTAACAGAGTCTACAAAAGCATATGTAAAGATTCTACAGTATG 6019  
Db 1613 GTGTGTGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAA 1672  
Qy 6020 CTAAGAGAAAGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6079  
Db 1673 GTGTAGCTGAAAAGTGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1732  
Qy 6080 AAGTATTCAAAATCTTATATTAATTAAGAAAGTGTATTAACAAAAGAAAGAAATTAAGTA 6139  
Db 1733 ATGTGAAGAAAGTGTAGCT-----GAAAATGTTGAAGAAATGCTAGCTCCAACTGTG 1786  
Qy 6140 AAAAAGAGTTTGTATCCAAATAGTTCAGTACTCATATCTTTAAATCTTTATTTGCGAAATG 6199  
Db 1787 AAGAAATGTAGCTCCAACTGTTGAAGAAATTTGTAGCTCCAAAGTGTGTA---GAAAGTG 1843  
Qy 6200 CCGCTGTTTCAAGACAAAGCCGAGAGTGTGACAGAACTGTTTAATTAATCAAGAGTTTATGAG 6259  
Db 1844 TGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAAGTGTGAAGAAAGTGTAGCTGAAAATG 1903

Qy 6260 AAACAGAGCTCTGTAGAAAATTTCTATTAATTAATGCAAAAATTAATTTCTGTAAATTCAG 6319  
Db 1904 TTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATCG 1963  
Qy 6320 GAGATTACAGAAATTCATTCGAGTAGTAGTCTGTTGTGTGTGTGGAATGTAGAG 6379  
Db 1964 TAGCTCCAACTGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTAGCTCCAACTG 2023  
Qy 6380 TAGAGCTCTTCTGATACCAATTAATTAATAAG---AAATACCAAGCAAGAGTGTGA 6436  
Db 2024 TTGTGAAGAGTGTGCTCCAACTGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAATG 2083  
Qy 6437 AAATACCAATGCTGATGAAGAGTTTCGAGAGAAAGAGTGAATTTACAGCAGATTCTAAGC 6496  
Db 2084 TAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATG 2143  
Qy 6497 AAGGAATTTCTCTTTTGTGAGTGTGAGAGTGTGAGAGCCGAGGTGAGAGCCGAGTGTGAG 6556  
Db 2144 TTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATGTAGCTCCAACTGTTGAAGAAATCG 2203  
Qy 6557 GAACCGTTTCCGTAAATCAATTTGACAGAAAGAGAGAGTGTGTAAGAGCAAGCA 6616  
Db 2204 TAGCTCCAACTGTGAAGAAATGTAGCTCCAACTGTTGAAGAAAGTGTAGCTCCAACTG 2263  
Qy 6617 TTTTGTGAAGAAAGCTGAGATTACAGCAAAACGTTATAGTCTGTTGCAATTTGGAATG 6676  
Db 2264 TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATG 2323  
Qy 6677 CCGAGTGTGAGTGTGCTGCAAAAGAGCTGGAATTTGAGACAGCAGTGTGCAATTCAGAA 6736  
Db 2324 TAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAACTG 2383  
Qy 6737 ATGAATCAAAACAGAGCAAGAGTGAAGAAATTTCTAAATTT---ATGACTGAAACAGT 6793  
Db 2384 TTGAAGAAATTTGTAGCTCCAACTGTTGAAGAAAGTGTAGCTCCAACTGTTGAAGAAATG 2443  
Qy 6794 TAGATGTAAATGACAGAAATAGATTAATTCAGATCTGGAATCGTTCAACCGAGCTG 6853  
Db 2444 TTGCTGAAAACGTTGCAACAAATTTATCAGCAATCTTTAAAGTAATTTATTAAGTGTGTA 2503  
Qy 6854 GAATTTCTGACGCGGAGTATCTGAGAGTGTGTTCTGTCAATTAATTTGCAATTAAGTATG 6913  
Db 2504 TCGAACTGTAGGAATTAAGACAGTATATTAATTAAGATGAAGAAAGTGAAGAAATG 2563  
Qy 6914 AAACAGATGATCATAGTACTTTACACTTTCTAGATGATTAATGTAAGCTCTTA 6973  
Db 2564 TAGTACCAATATCTAGAAAACGTAGAGAACTACAGCTGAAAGTGTAACTACTTTTA 2623  
Qy 6974 ATAAATTTTGAATTCCTTTCAGACGCGTGTGAGAGCGCAGTCTTTCAGACAGTTACCG 7033  
Db 2624 GTAACTATTAAGAGAGATACAGAAAATCTATTTACTTAATGATATCTATTAAGAGAAAAT 2683  
Qy 7034 GAGTGTGTTCTGTTAACATTAATTAATGTTCTGTAGTACTCGAGTTTCAATTAATCTG 7093  
Db 2684 TAGAAGAACTCCAGAAAATGTATTAATGTCGCTTTTGAAGAAATACCAAGGTGAAGAG 2743  
Qy 7094 ATTGACTTCGCTGCGAG---AAAGTAAATGTAACGCGAAAGAGGAAAAAATTA 7150  
Db 2744 AAAAGAAAGAGTATAGATTAATTAAGAAAGTGAAGAAAGAGTGTGCTACCTTTAA 2803  
Qy 7151 AGCAAGACGACAAATGACAGAAATCGAGAGACAGCAATCGAGCCAAATGCTTGTGTA 7210  
Db 2804 TAGAAGCTGTGGAACAGAGCAAGAAAGAGCGCAATTAATTAACGGAATTAATTTGAAA 2863  
Qy 7211 ATAAATTTGGAACAGCTGTGAGAGATGAAGAAATTTGGAAGAAAGAAACAGAGTTT 7270  
Db 2864 ATTTAGAAAGAAATGACAGTGAAGATGAAGAAATGTTGTCAGAGAAATTTAGAGAAATTA 2923  
Qy 7271 TAAAACTTTAGACGAAGTTTAACAAAGAACAGATTAATAAGATGCTACGAAA 7330  
Db 2924 ACGAACTGTATTTTAATCTGTATTAATTAAGTAAGAGAAACAGTAAGAAATTTAGCGAG 2983



```

; PRIOR APPLICATION NUMBER: PCT/FR96/00894
; PRIOR FILING DATE: 1996-06-12
; PRIOR APPLICATION NUMBER: FR 95/07007
; PRIOR FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 6152
; TYPE: DNA
; ORGANISM: P. falciparum
US-09-742-096-1

Query Match      1.2%; Score 116.2; DB 3; Length 6152;
Beet Local Similarity 41.7%; Pred. No. 9,6e-10;
Matches 1747; Conservative 0; Mismatches 2373; Indels 67; Gaps 14;

QY 5126 GTTCTGATGAAACGGAAGCTTTAGTTAGTATTCGAGTTTGAAGAGTAAGTCTTTCA 5185
DB 1017 GTGTGAAGAAAGTAGCTTCAAGTGTGATGAAAGTATAGATTCAAGTATTGGAAGAA 1076
QY 5186 ATGTAGATGCAAAAGATCAAAAACATTAATACAA---TTGCCGAAATGCAAAATGGAG 5242
DB 1077 ATGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAAAGTGTGTAGAAAAGTGGCTCAA 1136
QY 5243 GAAAAACGGCTGGAAGTTGAGACACAGTTGCTCATACAAATTGGAACAAATCAGTTA 5302
DB 1137 GTGTGAAGAAAGTAGAAGAAATGTTGAAGAAAGTAGTAGTAAATGTTGAAGAA 1196
QY 5303 TAGCTATTGTAAAAACAGTAAATTTACACGGCGAATGATCAAGATGAAAAATATCA 5362
DB 1197 GTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA 1256
QY 5363 ATGTAGCTGCAAAAGATTATATGACCAATATATAGCAGTCCGAGTTGAGAGGCA 5422
DB 1257 ATGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAA 1316
QY 5423 AAGGAGCTCTGTGCAAGAGAGCTTGTGCAAGTACTACCTGTAATAAGACAGTTTCTGTC 5482
DB 1317 TTGTAGCTCCAAAGTGTGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA 1376
QY 5483 ATGTGATCAAACTGATATTGCAAAAGATTAGAGAGAAATTAAGAAATTAAGAA 5542
DB 1377 ATGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1436
QY 5543 AAGCAATGTTATGTTCTAGCTGAAATATACAGTCAAGTGTGTCACAAATGCGACAGTGC 5602
DB 1437 GTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA 1496
QY 5603 TTTCCGAGCAAGTGAACAGCTGTGACAGTGTGAGTAGCAGTTAATAAATTACAC 5662
DB 1497 ATGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAA 1556
QY 5663 AAAAATCTTGCACATATTAATAATAGTACTCAAAATGTACGAATGCTTTGTGTAATA 5722
DB 1557 TTGTAGCTCCAAAGTGTGTGAAGAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA 1616
QY 5723 GC---AAATCTCATCTATTATAAACAATTGGAATTGAGCTGAGAGTGGAG 5779
DB 1617 ATGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1676
QY 5780 GAGCTGAGATGACAGATTCTGTAGCAGTGAATAGATTGTAAATATATACATAGACAGAT 5839
DB 1677 GTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA 1736
QY 5840 TAAATCATGCAAAATCACTGCGAAGGAATGTGAGATTATTAACAAGTCTGATGCGG 5899
DB 1737 ATGTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAA 1796
QY 5900 TAAATGCTAATATATGACAGGACAGTGTCTGAGAGTGGCCCGTCACACATAGAGCTCAA 5959
DB 1797 TGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATGTTAGCTCAA 1856
QY 5960 CCAGTGTGATGAATTAACAGATCTACAAAGCATATGTAAAGATTCTACAGTGAATTG 6019
```

```

DB 1857 GTGTGTGAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAA 1916
QY 6020 CTTAAGAAAGAACAGATGATTAATTAATCTACTCAAGGCGACAGTGAATAAGTGTAGATA 6079
DB 1917 GTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAA 1976
QY 6080 AAGTATTCAAAAATCTTAATATTAATAACGAGACTTATCAAAAAAAGAAAAATTAATAATA 6139
DB 1977 ATGTGAAGAAAGTGTAGCT-----GAAATGTTGAAGAAATGTTAGCTCCAACTGTG 2030
QY 6140 AAAAAGATTTGTACCAATAGTTCAGTACTCATTAATTAAATCTTTATTGGCAATG 6199
DB 2031 AAGAAATGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAAAGTGTGTA---GAAAGTG 2087
QY 6200 CCGCTGTTCAAGACAGCGGAGTGTGAGGACATGTTAATATCAACAAGTTTATGAGAG 6259
DB 2088 TGGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATG 2147
QY 6260 AAACAGAAAGCTCTGTAGAAATTTCTATATTAATGCAAAACATTATCTGTAAATCG 6319
DB 2148 TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATCG 2207
QY 6320 GAGATTACAGAAATTCATCGAAGTAGTGGTTCTGTTGTGTGTGAGAAATGAGAG 6379
DB 2208 TAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGTTAGCTCCAAAGTG 2267
QY 6380 TAGAGCTTCTTGTGATACCAATTAATAAAG---AAATACAAAGCAAGTTGGA 6436
DB 2268 TTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAATG 2327
QY 6437 AAATCAATATGCTGATGAAGAGTTTCGAGAGAAAGAGCTGAATTTACAGCAGATTCTAAGC 6496
DB 2328 TAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATG 2387
QY 6497 AAGAAATTTCTCTTTGAGAGTCCGAGTCCGAGACCCGGGATGAGACCCGAGTGGCAG 6556
DB 2388 TTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATGTTAGCTCCAACTGTTGAAGAAATCG 2447
QY 6557 GAACCGTTCCGTTAAATCAATTTGCAAGAAAGACGAAAGTGTGTAAGAAAGCAAGAA 6616
DB 2448 TAGCTCCAACTGTTGAAGAAATGTTAGCTCCAAAGTGTGAAGAAAGTGTGCTCCAAAGTG 2507
QY 6617 TTTGTGTAAAAAAGCTGAGATTACAGCAAAACGTTATAGTCTGTTCAATTGGAATG 6676
DB 2508 TTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATG 2567
QY 6677 CCGAGTCCGAGTGTGCAAAAGAGAGCTGGAATTTGAGACAGAGTGGCAGTTACCAAG 6736
DB 2568 TAGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAAGTGTAGCTCCAACTG 2627
QY 6737 ATGAATCAAAACAGAGCAAGAGTGAATAATTGTAATAAT---ATGACTGAAACAAGT 6793
DB 2628 TTGAAGAAATGTTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAATG 2687
QY 6794 TAGATGTATATGCAAGAAATAGATAAATCAGGTAATCGAATCGGTTCAAGCGGAGCTG 6853
DB 2688 TTGCTGAAAGCTTGTGCAAAATTTATACAGCAATCTTTTAAGTAATTTAAGTGTGTA 2747
QY 6914 AAAACAGATATGATCATATGTAATCTTACCTTTCTAGTATGTAAATGTAAGTCTTA 6973
DB 2808 TAGTCACCAATATCTAGAAACGTAAGAAACGTACAGCTGAAAGTGTAACTACTTTTA 2867
QY 6974 ATAAATTTGCAATTCCTTGAACAGCCGGTGAAGAGCCGAGCTTGCAGACAGTTACG 7033
DB 2868 GTAAATTTTGAAGAGATACAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2927
QY 7034 GAGTGTGTTCTGTTAACAATTAATAAGTGTGTGATAGCTGAGTTCAATAATCTGTG 7093
```

Db 2928 TAGAAGAACTCCAGCAAAAATGATTAAGTCCCGCTTGAAGAAATACCAAGATGAAGG 2987  
Qy 7094 ATTTGACTTCGTCAGAGA---AAAAGTAATGTAAACGGCAAAAGAGAAAAATATTA 7150  
Db 2988 AAAAGAAAAGATATAGTGTAAATTGAAGAAAGTAAAGAGGTCGCTACACCTTTAA 3047  
Qy 7151 AGCAAAACGACGAATTCGAGAGTCGAGAGCGCAATCGGACCAATGCTTGTTA 7210  
Db 3048 TAGAAGACTGTGAACGAGGAGAAAGAGCGCAAAATCAATTACGGAATATTTGAAA 3107  
Qy 7211 ATAAATTTGGAACAGCTGTGAAGATAGAAAAATTTCTGAAAGAAAAGCAAGATT 7270  
Db 3108 ATTTAGAAAAGAAATCGAGTAAAGATTAAGAAATGTTCCAGAAATTTAGAAATTTA 3167  
Qy 7271 TAAAACTTTAGACGAAGTTAACAAGAACAGATAAAAAGTAAATGATGCTACGAAA 7330  
Db 3168 ACGAACTGTATTTAATCTGTATTAGATTAAGATAGAGAAACAGTAGAAATTTACCGGAG 3227  
Qy 7331 AAATCTTACATCAGCAGATATTTCTACGAAGATCTTCTGTAAGCGGATAGAGAG 7390  
Db 3228 AAAGTTTGAAGAAACAATGAATGATTAAGCATTTTATGTAATAATTTGATTAATGTA 3287  
Qy 7391 AATACGAGGAGAGGAATTAAGCATGTGAAAGCTTCTGATATTTATTTGAAAAAATG 7450  
Db 3288 AAGGAATACAGAAAAATTTATTAACAGTATG-----TTGGAATATGAACCACTA 3341  
Qy 7451 TAGATATTTCAACAGAGACAAGAAATATATATCTTCTAATGTTGGAACTGCAG 7510  
Db 3342 TAGATATCAATCAGAGAAAGGTTGATTTGATGAAATGATGCTGATGCTGATTTAG 3401  
Qy 7511 GTCTGCTCCGATCAGAGACGTGCGCTTACAAATTTAAAGAAATTCGGAATT 7570  
Db 3402 ATATAATAGAAAAATATGAAGAAAGGTTTATTAATTAATTAAGAAATTTCAAGTACG 3461  
Qy 7571 CTGTGAAAAATCTTTGTGAAGACGCTGAAAGAAAGTAAATGTTAGATGCGATAT----- 7625  
Db 3462 AAGGTTTCAAGAACTGTAACTGAACATGTAGAACAAATATGTAATGTGATGTTGATG 3521  
Qy 7625 ---TACAGAAATGTTGCTTTAACACATATCAAGCTCTGTAGAGCATTTGGAAATG 7681  
Db 3522 TTCCGCTATGAAGATCAATTTTGAATATTAATGAGGAGAGGTTGAAGAAA 3581  
Qy 7682 GAGCTGCTATGCAAAATTAATTTCTAAATGGAAGATCAATATCAGATATTAATAATTTCTA 7741  
Db 3582 TGTATTTTATTTGGAAGATGTAATTTAAAGTGAAGTGAATGTAATCTGTAGAGAAA 3641  
Qy 7742 AGTATATAGAAAAAATATTTGATGTTATGTTAAAGATTAATCGAAATGAGCGGAG 7801  
Db 3642 TTAAAGATGAACCGTTCAAAAAGAGTGAAGAAAGAACTGTTAGTATTTGAAGAAA 3701  
Qy 7802 CAAAAGATTTAACCGTAGAGCGGTAGCTGCCGAGCCATTAATCTCAAAAGCAAAAGATG 7861  
Db 3702 TGGAAABAAAATATTTGATGATGTTATGAGAGAGAAAAAGATTTAAACGACAAATGA 3761  
Qy 7862 AAATGATTCAGAGTTGAAATGAGAAAGATATTTTCAATGAAGAAAATAGATTAATCTA 7921  
Db 3762 TAGATCAGTAGAAGATCAATGAAATATCTTCAATTTCAAAAGAAATCGAATCTTA 3821  
Qy 7922 GCCCTTCTAAAGAAATTTGAAGAAATCAATCTCAAGTGAAGAAAAGAAACAGATGA 7981  
Db 3822 TTAAAGATTAAGAAAAAGATGTTTCACTAGTGTGAGAGGTTCAAGCAATGATATG 3881  
Qy 7982 CTGCTGAATCTCAAGAGCT-----TCTGTAGAGCAGTATGACGAGGAG 8027  
Db 3882 ATGAAAGTGTGAGAAAGTTTAAATGAAATATGAAAGAGAGATTAATGAAGATG 3941  
Qy 8028 AATTAATTTCCGAACAAAGATCCGAGACTCTTAATTTGAAAGTTAGTACAAATCCG 8087  
Db 3942 CTGTTGAATTAATGATTAATAGCAAACTTAATTTGAAGAACTCAAGAGTTTAAATGAG 4001  
Qy 8088 AAGAGATTTTCAAGCAATATATGATATGATATGAGCAACATTAATAATGAAGTAAC 8147  
Db 4002 TAGAAGCAATTTAATTAAGATGATGAGAAAAATTAAGAAATTAAGAAAGCC--ATTATC 4059

Qy 8148 AGCAATTTCTAAGCAGTAACAGGTTCTGTATTTGGAGAGATGGAGTCAACCAAGCAGA 8207  
Db 4060 AGAAGATTTCTAAGCAATTAATGATGCAAAAGATATCAATTAAGAAAAATTTGAAGA 4119  
Qy 8208 AGTACTGCTCAGGTAAACATATGTTAGAAAGTTGAGAGAGAAATTTGTTCAAGACAA 8267  
Db 4120 GGAACATGATTAATGACAGCGTTGATGAAAGTTGATTAATAAGATGTCGAAGAGA 4179  
Qy 8268 TCGATTGAATCAATTTCTAAAGTGAAGGTTTGATGAAGTAAAGTAACTGCTAAATC 8327  
Db 4180 CAAGATCGAAAAAGATCTGATTTAAAGATCTTGAAAGATTAATTTAAAG----- 4232  
Qy 8328 TTCTGTATATCAGGAATGAGAGAGAAATTCGCGAGCAGAGATGAATTAATCTACAGC 8387  
Db 4233 --AGTAAAGAAATCAAGAACTTGAAAGTGAATTTTAAAGATTTTAAAGATTAA 4290  
Qy 8388 ACAAAGTAATCTGAATCCGATGTTTACGAAGCAAGTTATGAAGATTAATGATTA 8447  
Db 4291 AACTATTGAACAGATATTTTAAAGAGAGAAAAAGAAATTAAGAAATCAATTTGAAA 4350  
Qy 8448 CACAAAAAATATATTTCAAGAGTCAATGCTTCTGTTAAATGATACAAAGATGAAC 8507  
Db 4351 ATTCGAAGAGAGCTGAAGAAATTAAGATCTTGACAGATATATTTAAAGAGATATC 4410  
Qy 8508 GAATATGAATCTTACGCGTACCGGCTGATGCAACAAGAACAAACAAAGCAATTAAC 8567  
Db 4411 TTCAATTAAGTTGAAGAGAAAAAATTAAGAAAGTACGAATTTAAAGAAAGAGT 4470  
Qy 8568 GAGATCAACAAGTTAATCTTCACTCACTGAATGAGAGAAACGTATCTCACTTCGTC 8627  
Db 4471 AGAATATTAATTAAGTGTATGCTGATGCCATTAATTAAGTTTGAAGAAAGATGATTAAGAA 4530  
Qy 4531 AGTAGATGA-TTAAAGAGAGATATATTAAGATGATTAAGGAGAGATATGA--ATTAG 4587  
Db 8628 AAAAGCTTGGCTAAAAATGAATATGAAATGAAATGAAAGAACTGAGAGCCTTACT 8687  
Qy 8688 CGAGCGGAAACAGACCGCTTGAATTAATCAAGAGTACTACAGAGCTTGGTTC 8747  
Db 4588 GGATATGATTAAGAAAGTTTGAAGATGTAACAACAACTTGAGAAAGAGTTGAATC 4647  
Qy 8748 AGGAATTTGGAAATTTGAGATTAATTTGAACAGATTCGACAGAGATTAATCAATTTGA 8807  
Db 4648 CTTAAAGATGTTTATCTAGTCAATTAAGCATGATGAAGAACAAATGAAGAAACAAGAA 4707  
Qy 8808 ACTCAACGAGACGAAACCAAGAGAGCTTGTGGAAGAAATGATTTCTGTGAAAA 8867  
Db 4708 AAAAGCTCAAGACCTAAGTTGGAAGATTAATTTAAAGAAAGGTTAAAGAAAGACC 4767  
Qy 8868 TACAATTTGAGGGAACAAATCACTCAATGAAGTAAAGCAGAAATTTGGAACCG 8927  
Db 4768 AAAAGAAAAAATTAACAAGAAAGAAAGTGAATTTGATTAATGAATTAAGAAACAAAGAA 4827  
Qy 8928 AAGTGAATATGATGCTTGAATGAACCTTGAATGATCTTAAGAGAAAAAGTGTG 8987  
Db 4828 TGAATATGTAAGGTTGAATGAATGAATGAATTAAGAAAGATGTAAGAAAGATAT 4887  
Qy 4888 AGAAGAGATTAAGAAAGATTAAGTTGAATTAAGTGAAGATTAAGTGAAGATAT 4947  
Db 9048 AGCCAAATCGAAGACATGTAATGTAACCTACTGCAAAACAAAGATATCAAGATTT 9107  
Qy 4948 AGGTGAAGCAAGAGTGAATTAATTAATGATCTCAAAAGAAAGAAACGATTTGAAAA 5007  
Db 9108 TACAAGCAAAAGTAAATATTTCTGGAAGAGAGCGCTGACGCTGACGTAATATC 9167  
Qy 5008 GGTAAAGCAAGAAAGAAATTTGAAGAAAGTTGAAGAGGTTGATGCTTTAA 5067  
Db 9168 GAATGATCAATTTCCAAATGATGATTAATTAATTTGCAAGACGATATGATCTTC 9227  
Qy 5068 AAAACAGTGAAGAGATTAATGAATATGTTCAAAAAATTTGATTAAGAAAGTGAATGAAG 5127

```
QY 9228 TCAATTAAATACCAAAATTCAAAATAATATTAATTACTTAGCATCA 9274
DB 5128 AGTATCTAAAGCTTAGAATCAAAAATGATGTACTAATGTTTAA 5174

RESULT 17
US-10-415-253-1
; Sequence 1, Application US/10415253
; Publication No. US20040067236A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Joe
; APPLICANT: Drulhe, Pierre
; TITLE OF INVENTION: Immunogenic Compositions Comprising
; TITLE OF INVENTION: Liver Stage Malarial Antigens
; FILE REFERENCE: B45250
; CURRENT APPLICATION NUMBER: US/10/415,253
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: PCT/EP01/12349
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: EP00203724.0
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5528
; TYPE: DNA
; ORGANISM: K1 Parasite Strain
US-10-415-253-1

Query Match 1.1%; Score 107.6; DB 7; Length 5528;
Best Local Similarity 41.6%; Pred. No. 2.9e-08;
Matches 1741; Conservative 0; Mismatches 2379; Indels 66; Gaps 14;

QY 5126 GTTCGATGAAGCGAAGCTTAGTATGATTCGAGTTGAAGAGTAAGTCTTTCA 5185
DB 941 GTTTGAAGAAAGTATAGCTTCAAGTGTGATGAAGATTAATTCAGTATTGAAGAA 1000
QY 5186 ATGTGATGCAAAAATCAAAAACAATTAATACAA---TTGCCGAAATGCAATGAG 5242
DB 1001 ATGTAGCTCAACTGTGAAGAAATCGTAGCTCCAAAGTCTTGAAGAAATGCGCTCAA 1060
QY 5243 GAAAACGGCTGAGAGTTGAGCAACAGTTGCTCATACAAATATGGAAGAAACATCAGTTA 5302
DB 1061 GTGTGAAGAAAGTGTAGAAAGAAATGTGAAGAAAGTGTAGCTAAATGTTGAAGAA 1120
QY 5303 TAGCTATTGTAAAAACAGTAATTTACACCGCGAATGATCAAGATGAAAAATATCA 5362
DB 1121 GTGTAGCTGAATAATGTTGAAGAAAGTGTGAAGAAATGTTGAAGAAATGTTGCTGAAA 1180
QY 5363 ATGTGACTGCAAAAATTAATTAATGATGACCAATTAATAGCAGTCCGAGTTGAGAGCA 5422
DB 1181 ATGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAA 1240
QY 5423 AAGAGAGCTGTGTGCAAGAGCTTGTGCAAGTACTACTGTAATAGACAGTTTCTTC 5482
DB 1241 TTGTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA 1300
QY 5483 ATGTGATCAAACTGATATTGCAAGATTTAGAGAGAAAGAAATATGAAATTAAGAA 5542
DB 1301 ATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAA 1360
QY 5543 AGGCAATGTTATGTTCTAGCTGAAAATACGAGTCAAGTGTGCACAAT--GCAGACGT 5600
DB 1361 GTGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTGCTGAAA 1420
QY 5601 GCTTTCCGAGCAAGTGAAGAGCTGAGTAGAGAGCTGAGTAGCAAGTTAATAATTAC 5660
DB 1421 TGTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAAT 1480
QY 5661 ACAAAATCTTGCAATATAAAAAATAGTACTCAAAATGTACGAAATGCTTTGTAA 5720
DB 1481 TGTAGCTCAAGTGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAA 1540
```

```
QY 5721 AAGCAATCTCATTCATCTATTAACAATTTGAAATTTGAGAGTGGAGTTGAGCTGAG 5780
DB 1541 TGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAG 1600
QY 5781 AGCTGAGTGCAGAGTTCTGTAGCAATGAATTAAGTTGTAATTAATACATAGCAAAATT 5840
DB 1601 TGTAGCTGAATAATGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1660
QY 5841 AAATCATGCAAAAATTCATGCGAAGGAAAATGTCGAGATTTACAGAGTCTGATGCGGT 5900
DB 1661 TGTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAT 1720
QY 5901 AATTGCTAATTAATGCAAGAACAGTGTGAGAGTGGCCGTCAGCAATAGAGCTCAAC 5960
DB 1721 CGTAGCTCAACTGTGTGAAGAAATCGTAGCTCCAACTGTGAAGAAATGTTAGCTCCAAAG 1780
QY 5961 CAGTGTGAATGAATTAACAGATCTCAAGAACATATGTAAAGTTCTACAGTATGTC 6020
DB 1781 TGTGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAAGTGTGAAGAAAG 1840
QY 6021 TAAAGAAAGAACAGATGATTAATTAATTAATCAAGGCAAGTAGTAAGTGTAGATTA 6080
DB 1841 TGTAGCTGAATAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAA 1900
QY 6081 AGTATTCAAAATCTTAATTAATTAACGAAGACTTATCACAAAAAGAAAAATAGTATTA 6140
DB 1901 TGTGAAGAAAGTGTAGCT-----GAAAATGTTGAAGAAATCGTAGCTCCAACTGTGA 1954
QY 6141 AAAAGATTTGTTACCAATAGTTCAGCTCACTCAATCTTAAATCTTTATTTAGGCAATGC 6200
DB 1955 AGAAATCGTAGCTCCAACTGTGAAGAAATGTTAGCTCCAAAGTGTGA---GAAAGGT 2011
QY 6201 CGCTGTTCAAGAACCGAGTGGCAGAACTGTTAATATCAACAAGTTTATGAGA 6260
DB 2012 GCTTCAGAGTTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTGAAGAAATGT 2071
QY 6261 AACGAGCTCTGTGAAGAAATCTAATTAATTAATGCAAAACATTAATCTGTAAATCAG 6320
DB 2072 TGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATGT 2131
QY 6321 AGATTAACGAATTCATCGAGATGATGAGTTCTGTGTGTGTGTGTGTGAATGTTAGAGT 6380
DB 2132 AGCTCAACTGTGTGAAGAAATCGTAGCTCCAACTGTGAAGAAATGTTAGCTCCAAAGT 2191
QY 6381 AGGAGCTTCTGTGATCAATTAATTAATTAAG---AAATCAAGAACAGAGTTGAAA 6437
DB 2192 TGTGAAGAAAGTGTGCTCCAAAGTGTGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGT 2251
QY 6438 AACTACATGTCGTGATGAAGGTTTCGAGAAAGCTGAAATTAACAGCAATTCATACCA 6497
DB 2252 AGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGT 2311
QY 6498 AGGAAATTTCTCTTTTGAAGTGTGAGTGTGACGACGAGCCGGGTAGAGCCGAGTGGCAG 6557
DB 2312 TGAAGAAAGTGTAGCTGAAAATGTTGAAGAAATCGTAGCTCCAACTGTTGAAGAAATGT 2371
QY 6558 AACCGTTTCGTAATCAATTTGCAAGAAAGCGAAGTAGATGTGAAGAAAGCAAGAT 6617
DB 2372 AGCTCCAACTGTGAAGAAATGTTAGCTCCAAAGTGTGTGAAGAAAGTGTGCTCCAAAGT 2431
QY 6618 TTTGTGAAGAAAGCTGAGATTAACAGCAAAAGTATAGTCTGTGTGCAATTTGAAATGC 6677
DB 2432 TGAAGAAAGTGTGAAGAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGT 2491
QY 6678 CGCAGTCCGAGTGTGCAAAAAGAGCTGGAATTTGAGACAGCAGTGGCAGTTACCAAGA 6737
DB 2492 AGCTGAAAATGTTGAAGAAAGTGTAGCTGAAAATGTTGAAGAAAGTGTAGCTCCAACTGT 2551
QY 6738 TGAATCAACAGAGAGAGTGAAGAAATTTCTAAATTT---ATGACTCGAAACAAGTT 6794
DB 2552 TGAAGAAATTTAGCTCCAAAGTGTGAAGAAAGTGTAGCTCCAAAGTGTGAAGAAAGTGT 2611
QY 6795 AGATGTATACAGAAATGAGATTAATCAAGTACTGAATCGGTTCAAGCCGAGAGCTGG 6854
```

D	b	2612	TGCGAAAAAGCTTGCAACAAATTTATCGAACAAATCTTTTAAGTAATTTATTAGGTGTAAT	2671
O	y	6855	AATCTTGCAAGCCGGAATATCTGGAGAGGTTTCTGTCAATAAATATTCGCAATTAAGGTAGA	6914
D	b	2672	CGAAACTGAGAAATTAAGGACAGTATATTTAATGATAGATGAAAGATTAAGAAATATGT	2731
O	y	6915	AACAGATATCGATCATAGTACTTTACACTCTTCTACTGATGTAAATGTAAAGCTTTAA	6974
D	b	2732	AGTCACCAACAATCTAGAAAAAGTAAAGAAACTACAGCTGAAGTGTAACTACTTTAG	2791
O	y	6975	TAAATTTTCGAATTCCTTGACAGCCGGTGGAGAGCCGCAAGTCTTGACAGTACCG	7034
D	b	2792	TAAATATTGAGAGATACAAAGAAAAATCTATTACTTAATGAATCTATTAAGAAAAAT	2851
O	y	7035	AGTGTCTTGTTTACCTAATAATAGTCTGTGATAGCTCGAGTTCACAATACTCTGA	7094
D	b	2852	AGAAAGATCCACGAAATGTATTAAAGTCCGCTTGTAAGAAATATCCAAAGTGAAGAGA	2911
O	y	7095	TTTGACTTCGCTACGAGA---AAAAGTAAATGTAAACGGCAAAAGAGAAAAATATTA	7151
D	b	2912	AAAGAAAGAAATATAGTATGTAAATTTGAAGAAAGTAAAGAAAGAGGTGCTACACTTAAT	2971
O	y	7152	GCAAAACGACGAATAATGACGAATCGAGAGACGCAATGGAGCCAAATGCTTTGGTAA	7211
D	b	2972	AGAAACTGTGGAAACGAGCAGAAAGAAAGCGCAATATCAATTAACGAAATATTTGAAA	3031
O	y	7212	TAAATTTGGAAACACCTGTGAGAAATAGAAAAAATTCGAAAGAAAGAAACAGAGTTT	7271
D	b	3032	TTTAGAAGAAATACAGTAGAAAGTAAATGAAAAATGTTGACAGAAATTTAGAAATTTAA	3091
O	y	7272	AAAACTTTAGACGAAGTTAACAAGAAACAGATTAATAAGTAATGATGCTACGAAAA	7331
D	b	3092	CGAAACTGTATTATTAATCTGATTAATAGTAAGTAAGTAAGAAACAGTAAATTTGCGAGA	3151
O	y	7332	AATCTTAACATCAGCAGTATTTCTACAGAAAGTACTTCTGTAAAGCGATAGAGAGA	7391
D	b	3152	AAGTTGAAAAACAATGAATATGATTAAGCAATTTTGTAGTAAATATTTGATATGTAAA	3211
O	y	7392	TACTCAGGGAAGAAAGATTAAAGCATTGTGAACCTTGATTTATTTGAAAAAATGT	7451
D	b	3212	AGGAATACAGAAAAATTTATTAACAGGTAT- - - - -TTTGAAGTATAGAAACAGAT	3285
O	y	7452	AGATATTACAACAGACAGAATTAATATCACTTCTACTGCTGTTGGAACTGCAG	7511
D	b	3286	AGTAACTCAATCAGAAAGAAAGTTGATTTGAATGAAATATGTTAGTTCGATTTTGA	3325
O	y	7512	TCTTGCTTCGCAATCAGGACAGTGGCAGTTACAATTTTAAAGAAATTCGAGTTAC	7571
D	b	3326	TAAATTAAGAAATAATGAAGAAAGGTTTATTAATTAATTAAGAAATATTTCAATACTGA	3385
O	y	7572	TGTTGAATAATCTTTTGTGAAGACAGCTGAAAGAAAGTAAATGTAGATCGATAT- - - - -	7625
D	b	3386	AGGTGTTCAAGAACTGTACTGAAACAGTGAACAAATATATATGTGATGTTGATGT	3445
O	y	7626	---TACAGGAATGTTGCTTTAAACGCAATCAAGGCTCTGTAGAGCAATTGGGAATAG	7682
D	b	3446	TCTGTCTATGAAGATCAATTTTTRGGAATTTAATATAGCAGAGAGGTTGAAAGAAAT	3505
O	y	7683	AGCTGCTATGAGAAATTAATTTCTAATGAAAGTCAATATCAGTATTAATAATTTCTAA	7742
D	b	3506	GTTTTTTATTTTGAAGATGTATTTAAAGTAAAGTATGATTAATCTGTAGAAAGAAAT	3565
O	y	7743	GCTATTAGAAAAAATATTTGATGTTATTTGTAAAGATTAATTCGAAATTTGAAGCGGAAC	7802
D	b	3566	TAAAGATGAACGGTTCAAAAAGAGTGAAGAAAGAAACGTGTAGTATTTATTTGAAGAAAT	3625
O	y	7803	AAAAGATTAACCGTAGAGCGGTAGCTGCAGGACCAATTAATCTCAAAAGCAAAAGATGA	7862
D	b	3626	GGAAAGAAATATTTGATGATGTTTATGAGAGAAAGAAAGATTTTAAACGACAGATGAT	3685
O	y	7863	AATGAATTCAGAGTTGAATTTGAGAAAGTATTTTCAATGAAGAAATAGAGTAACTAG	7922

Db	3686	AGATGCGATGAGAAATCCATAGAAAATCTTCAGATTCTAAAGAAAGAACTGATCTAT	3745
Qy	7923	CCCTTTCAAAAGAAATTGAAAGAGAAATCAATGTCAAAGTGGAAAAAGAAAACAGAGTCAC	7982
Db	3746	TAAAGATTAAGAAATAAAGATGTTTCCACTGATGTTGGAAAGATTCAACACATGATATGGA	3805
Qy	7993	TGCTGAATCTCAAGAGCT-----TCTGTAGAGACGATGACGGGGCCGGA	8028
Db	3806	TGAAAGTGTGAGAAAGTTTTCAGATTGAAAAATATGGAAGAGAGTTTAAATGAAGATGC	3865
Qy	8029	ATTATTTCCGAAGCAAAAGATGCCGAAGCTCTTATTTGAAAGTTACTACAAATCCGGA	8088
Db	3866	TGTTGAAATTAATATCACTTACTACCAACTATTTGAAGAACTCAAGATTTAAATGAAGT	3925
Qy	8089	AGAAGTATTTTTCATGACAGATTAATGTGAATATGGAACCAACATTAATGAAGTTACA	8148
Db	3926	AGAGCAGATTTAATTAATAAGATATGAAAAATTTAAAGAAATTAGAAAAGC--ATTATCA	3983
Qy	8149	GCAATTTTCAAGACATTAACAGTTCGTATTTGGAGAGAGTTGGAGTCACCAAGCGCAGAA	8208
Db	3984	GAAAGTCTTAAAGAAATTAATGATGCAAAAGATATACATTAGAAAAAGTTATTTGAAGAG	4043
Qy	8209	GCTACTGCTGACGGTAAATTAATGATGTAAGAGTTGAGAGCAAGAAATTTGTCAGAAACAAT	8268
Db	4044	GAACATGATTAATACGACGACGCTTGATGAACTTTGAAATTTAAAGATGTCGAAGAAAGAC	4103
Qy	8269	CGATTGAATGCAATTTCTTAAAGTAGAAAGTTTGATGAAGATAAAGTAACTGCTAAATCT	8328
Db	4104	AAGATCGAAAAAGATATCGATTTAAAGATCTTGAAAGATATATTAAAAAG-----	4155
Qy	8329	TCTGTAGTATCAGGAATATGAGAGAGAAATGCCGACACGAGATGAATCTTCTACAGCA	8388
Db	4156	-AAGTAAAGAAATCAAGAACTTGAAAGTAAATTTTAGAAGATTATTAAGAAATTAAAA	4214
Qy	8389	CAAGTAATACGATCCGTGTTCCGTTACGAAGCAAGATTATGAAAAATTAATGATTAC	8448
Db	4215	ACTATTGAAACGATATTTTAGAAGAGAAAAAGAAATAGAAAAGATCATTTTGGAAAA	4274
Qy	8449	ACAAAAAATATATTTTCAGAAATCAATGCTCTTCTTAAATGATACAAAGATGAAGCG	8508
Db	4275	TTTCGAAGAAAGCGAAGAAATTAAGATCTTGAAAGCATATATTAAAAAGATATCT	4334
Qy	8509	AATATAGAATCTTTAGCCGTGCCGTGTCAATGCACAGGAACAAACAAAGCATTTACG	8568
Db	4335	TCATTAGAAAGTTGAAGAAAAAAAATTTAGAAAGTACAGATTTAAAGAAAGGTA	4394
Qy	8569	AGATCAACCAAGTTAACTTTCACACCTGAATGAGAGAAACGTATCCAACTCGTCA	8628
Db	4395	GAACTATTAATTAAGTGATGATGCCATATTAAGGTTTGGAAAGATGATTTAGAGAA	4454
Qy	8629	AAAGCTTTGGCTAAATATGAAAAATTTATGGAATGTAAAGCAATCGAGAGCCTTACTC	8688
Db	4455	GTAAGATGA-TTTTAAAGAAAGATATTTAGACATGTTAAAGGAGATATGGA--ATTAGG	4511
Qy	8689	GGAGCGGAAACGACGCCGTGAAAAATTATACAAAGACTACTACAGAGCATTTGGTTGCA	8748
Db	4512	GATATGGAATTAAGGAAGTTTGAAGATTAACAACAAACTTGGAGAAAGGTTGAATCC	4571
Qy	8749	GGAAATTTGGGAAATTTGAGATTAATTAAGAACATTTGAAGAGATTAATACGATTTGTAAG	8808
Db	4572	TTTAAAGATGTTTATCTAGTGCAATTTAGCATGATGTAAGAACAAATGAAACAAAGAAA	4633
Qy	8809	GTCAACGAGACGAAACCAAGAGAGTCTGTGCGAAAGAAATGATATTTCTGTGAAAAAT	8866
Db	4632	AAAGCTCAAAAGACTTAATGTGGAAGAAATATTAATTAAGAAAGAGGTTTAAAGAAACCA	4691
Qy	8869	ACAATTTCAAGGGGAAACAAATCATCACTTGAAGATAAAGCCAGATTTGTTGGAACCGGA	8928
Db	4692	AAGAAAAAATAACAAAAAAGAAAGTAAGTTGATATTTAAGATTAAGAAACCAAAATAT	4751
Qy	8929	AGCTGTAATGTGATGCTTTGAATGAATCTGATGTGATGATTAACAAGAAAAAGTGTGGC	8988
Db	4752	GAAATGTAGAAAGTTGAATGAAGATGAAGATATAGAAAGAGATGTTGAAAGAGATATTA	4811

QY 8989 TATGTGGAAATGTGTAATGTAATTAATGATTAAGAAAATGTAGAA 9048  
DB 4812 GAAGAGATATAGAGAGAGATTAAGTTGAGATATAGATGAAGATATAGAGATATTA 4871  
QY 9049 GCCAAATTCGAGAGACATGCTATTTGAGAACTAGGAAAACAGAAATATCAAGCATTT 9108  
DB 4872 GGTGAAGACAAAGATGAAGTTATAGATTTAATGATCCAAAAGAGAAAACGATTTGAAAAG 4931  
QY 9109 ACAAGAGCAAAAGTAATATCTTGGAAAAGAGACGCTGACGCTGCAATATCG 9168  
DB 4932 GTTAAAGCGAAAAGAAAATTTAGAAAAAGTTGAAGAGAGTGTATGCTTTAAA 4991  
QY 9169 AATGACACATTTCCAAATGAGATGATATTAATAATTTGGCAAGACAGATGCACTTCT 9228  
DB 4992 AAACAGTAGACGAGATGATGAATATGTTCAAAAAATTTGATTAAGAAAGTTGATTAAGA 5051  
QY 9229 CAATTATATACCAAAAATTTCAAAAATTAATTAATTAATTAATTAATTAATTAATTA 9274  
DB 5052 GTATCTAAAGCTTTAGAAATCAAAAATGATGTTACTTAATGTTTAA 5097

RESULT 18  
US-10-668-749A-1  
; Sequence 1, Application US/10668749A  
; Publication No. US20040110205A1  
; GENERAL INFORMATION:  
; APPLICANT: Agilent Technologies  
; TITLE OF INVENTION: Methods and Systems for Nanopore Data Analysis  
; FILE REFERENCE: 50112-1580  
; CURRENT APPLICATION NUMBER: US/10/668,749A  
; CURRENT FILING DATE: 2003-09-23  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 1  
; LENGTH: 1300  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic construct  
US-10-668-749A-1

Query Match 0.8%; Score 80.4; DB 7; Length 1300;  
Best Local Similarity 41.4%; Pred. No. 0.0086;  
Matches 537; Conservative 0; Mismatches 761; Indels 0; Gaps 0;

QY 8426 AAGATTATGAAATTAATGTTTACAAAAAATTAATTTTCAAGTCAATGCTCTTGCTT 8485  
DB 1 AA 60  
QY 8486 TAAATGATCAAGAATGAAGCAATATATGATCTTTAGCGGTAGCGGTGTCATGCAC 8545  
DB 61 AA 120  
QY 8546 AAGACAACAACAAGCATTTAGAGATCAACAAGTTAACTTCTCAACTGTAAATGAG 8605  
DB 121 AA 180  
QY 8606 GAAACGTATCTCACTTCTGCAAAAGCTTTGGCTTAATAATGAATAATTTGAAATGTA 8665  
DB 181 AA 240  
QY 8666 AAGGAAGTGAAGAGCTTAGTCGAGCGGAAACAGCAGCCGTTGAAAATTTATCAAGA 8725  
DB 241 AA 300  
QY 8726 GTACTACAGAGCATTTGTTGCAAGAAATTTGGAAATTTGAGATTAATTTAGAACGATTG 8785  
DB 301 AA 360  
QY 8786 CAAGAGATATATCGATTGTAAGATCAACGAGACGAAACAAAGAGGCTTTGTCGAA 8845  
DB 361 AA 420

QY 8846 AGAATGATTTCTGTGAAAAATATCAATTTCAAGGGAACAACAAATCATCATTTGAGATA 8905  
DB 421 AA 480  
QY 8906 AAGCCAGATTGTTGAAACCGAAGTGAATTTGATGCTTTGATGATCACTTGATGAG 8965  
DB 481 AA 540  
QY 8966 ATCTACAGGAAGAAAGTGTGCTATGTTGGAATTTGATTTGGAATGTTGATGTAATA 9025  
DB 541 AA 600  
QY 9026 ATGTGATTAAGAAAAATTTAGAACGCCAAATTCGAGACATGCTATTTGAGAACTACTG 9085  
DB 601 AA 660  
QY 9086 GAAAAACAAGATATCAAGCATTTTCAAGAGCAAAAGTAATATTTCTTGGAAAAGAGACG 9145  
DB 661 AA 720  
QY 9146 CTGCAAGCTGCAAGCTGCAATATCGAATGTACATTTCCAAATGAGATGATATTAATAAT 9205  
DB 721 AA 780  
QY 9206 TGCCAAAGCAGTATGATCTTCTCAATTATATACCAAAAATTTCAAAAATTAATTAATTAAT 9265  
DB 781 AA 840  
QY 9266 TAGCATCAAGTAGTGAATCGAATGTAATGTTTCAATGGGTGCTGAAGCAAGAGTGCA 9325  
DB 841 AA 900  
QY 9326 GAGCCAAAGCAGCTGATGTTAAGATCAATTAATTAAGACTAATATGTTGATTTAG 9385  
DB 901 AA 960  
QY 9386 CAGCAAAAATTTAAACAGAGGAAACATCAATGTATATCCGATATGATTAATAATTAATA 9445  
DB 961 AA 1020  
QY 9446 ATATATAGTAAGCAAAATTTCAAGCTATTTGCGATGCCAAAGTCATCTGCACTGCTT 9505  
DB 1021 AA 1080  
QY 9506 CGGCACTGCCATATTTGAATAATGAATGTAATTTAATTAATGCGATCCGAGAAATTTA 9565  
DB 1081 AA 1140  
QY 9566 AAAATATCTGGCAAGATTTGAAAGGAAAGCTAATTAATAAAGCTCGGTAGATCTTAATC 9625  
DB 1141 AA 1200  
QY 9626 AAGTAGCTGTATACGATTAATATATATATATATATATATATATATATATATATATAT 9685  
DB 1201 AA 1260  
QY 9686 TGACATATCATCAAGAGAGAGAAAGGAAAAA 9723  
DB 1261 AA 1298

RESULT 19  
US-09-815-242-4761  
; Sequence 4761, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 4761  
LENGTH: 7434  
TYPE: DNA  
ORGANISM: Staphylococcus aureus  
US-09-815-242-4761

Query Match 0.8%; Score 79.2; DB 3; Length 7434;  
Best Local Similarity 41.3%; Pred. No. 0.0029;

Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

QY 3782 TTCAGATGAAACAAGTCTTTGGTTAAATAATAGTATTCGAAATGCCATTAATTTA 3841  
DB 3659 TTAATTAATAAAGCACTGCACTGATGAATGAAATCAGACATTTTAATTAACAATTGCAAG 3718  
QY 3842 AAGTAATGCTTTGAGTGAAGAACTCAAGTACAGACAGACAGAGTTTGAAGACATTA 3901  
DB 3719 ACATTCAGACTACGCCAGATGCAACGATGAAGAAAAACAAGCAGTGTGCAAGCA 3778  
QY 3902 AAGAAAGTGAAGCAAGAAAAAGTTATCTATTTGGAACTTCTGTTCTATCACTTGA 3961  
DB 3779 ATACTGAAATGATAAGCAATTCAGCACTTCACAGCACTATTAACGACAAGTTG 3838  
QY 3962 TGAACAATGATTTCTGCAAAATCAGAAATTAATCAGTAGCAGAGATCGAAAGCC 4021  
DB 3839 ATGAGACTAAAGCAATGCAAGACGCAATTAATCGGTACACCAAAAGTTGTGAAG 3898  
QY 4022 AAAAATGATGTTGATGTCACTGCTTATCAAGCCGACACCAAGTGAAGAGCTTTTA 4081  
DB 3899 AACAGCGCTAAGATGAATGATCAATTAACAAACGCAACAAATGTTATCAATTA 3958  
QY 4082 ATTTACAGCTGAAGTCAATGAACTGTAGGGCTACTGTGACTGTGCTTCCAAATTAA 4141  
DB 3959 ATGATCAGAACGCTACAAAGAAAGAAAAAGACACACTATTCAAATTAAGCAACGAG 4018  
QY 4142 ACAACAAGTAATGCTTCATTAATGATGAGGAGATATCTAACTTAATCGAGCGGAG 4201  
DB 4019 TTAACAGACCGAAAAATTAATTAATTAACGCTGCACTGATTAATGATGATCAGGCA 4078  
QY 4202 CAAGAAGCTTTTGAAGCACTCAAGTACTGCTCAGTGAAGCAGGAGGAGCAATTA 4261  
DB 4079 AAGAGCTGGAAGAAATTCATTAAGCAACGCAACAGGAGTTAATCAATTAAG 4138  
QY 4262 GTTCTGAGCGGATTAAGAAATTAATCAAGGGGCTGTTCTGCTCAATTAAGATTGAAG 4321  
DB 4139 CTAAATATGATGTTGATCAAGCTGTGCAACTCAAAATCAAGCAATTAATTAACAATG 4198  
QY 4322 AGGTGAAGCTAGCTGATTAATCTTCATCGAAGAGCTAATGAATCATGTCAATG 4381  
DB 4199 GTGCTACACTGAAGAAAAATGACGCAAAAGATTTAGTTTAAAGCTTAAGAAAAAG 4258

QY 4382 CCAAGATGCAAGAAAGTCTGATCTAGCAAAAAGATATCAGCTTTACTAATATGAA 4441  
DB 4259 CGTATCAAGATATCTTAATATGCAAAACACTAATGATGTTACGCAATTAAGATCAAG 4318  
QY 4442 AAGATAAAAATATTTAGAGATCGTGTATTAATACACTGGAATGTTATTAACGA 4501  
DB 4319 CAGTTGCGATATTCAGAGTATTACTGAGATACAAATTAAGATGTTG-----CGA 4372  
QY 4502 AGGAACAATGAAAAAGCAAAAGAAAAAGAGCGGCTCATTTGTAATGCTGTTAT 4561  
DB 4373 AAGATGAATTTGCAACAAGCAAAAGCAAAAGCCCTTAATTCACAAATGCAAGATG 4432  
QY 4562 CGGTTGCGAAGGATTAATCCGCTGAGAGTAGTATTCAGTCACTATGTTTAA 4621  
DB 4433 CGACTACTGAAGAAAAAGAACAGCAATTAACAGTGAAGCAGCAATTAACACAAGTA 4492  
QY 4493 ATCAAAATATTTGAATAATGCAAGTCAATCGATGATGTAACACTGCAAAAGATATGCA 4552  
QY 4682 CGAAACATGTAATGTGAAGCAAAATCATCTGTTGTTGTGTAATCGGCTTGTGAC 4741  
DB 4553 TTCAAGCAATTCACCAATTCAGAGCATCAAGATGTTAAACGAATCGAAGC-GGA 4611  
QY 4742 TTGCTATGAGCAAGATGCTTTTTCAGGATGCGATCTGAGCATGCGCAAGCTTATCA 4801  
DB 4612 TTGCTACTGAATCAAAATTAATACTGAATATCTTAATTAATGAGCT----- 4665  
QY 4802 ATGACAGATTCGAAGGTGATTAAGAAAGAAATTTCTGCTGATTCCTTAATGTAAGC 4861  
DB 4666 -----ACTTAATGAAGAAAGTAAGTAATGTAAGCAAGTATGACAGATGTAAGAG 4720  
QY 4862 CAATTAATTCATCTTGGGTAATGTTGCGGAAACCAATGCGGTTCTTTCTACGG 4921  
DB 4721 GTTTAAATTAATTAATGACAGCACTACAGTATGATTAATCTGCTTAAGATACAG 4780  
QY 4922 CGGTGAGAGCTTTTTCGAAATATCTCTTCAATTAATTAACCTGCTTTGATTAACG 4981  
DB 4781 CAGTCAAAAGTTCAACAACTTCATGCAAACTGTTAAGAACACAGAGTTAAAAAG 4840  
QY 4982 GAACGAAGTAATCTTTTATGTAAGAAATATCAAAAGTCAATGATACAGCTTTGAATG 5041  
DB 4841 AATTGATCAAGCTCAGCTGATTAAGAAACAAATTAAGAACAAACCAATGATCAAC 4900  
QY 5042 ATTCATATTAACAAGCTTCTGCTGAGCGCTGCAATTAATTAAGCGCTGGAATG 5101  
DB 4901 AACAAAGTAATTAATGATGCAAAACAAAGTTGATGATTAATTAATCAAGGAAACAA 4960  
QY 5102 GAGGAATGATCTGCTCAATCGTGTCTGATGAACGGAACCTTAATGATGATTCG 5161  
DB 4961 ATGTGATCAATCAATCAACAAATGAATGTTGATTAATGCACTTAAGAGAAAGCTTA 5020  
QY 5162 AGTTGAAGAGTAAGTTCTTCAATGTGATGCAAAAGATCAAAACAAATTAATACAA 5221  
DB 5021 AATTTATGCAATTAACATTTAGTGAATCAAAAGATGCTTTAGTAAATTTGAAG 5080  
QY 5222 TTGCGGAATGCAATGAGAAAAAGCGCTGAGGTTGAGCAACAGTTGCTCATCA 5281  
DB 5081 ATGCAATATTAATCTTAAGTAAGAAAGCGGATTAATCTTAACCATGCACTTAAGTGA 5140  
QY 5282 ATATTGAAAAATCAATCAATTAATGATTAATTAATTAATTAATTAATTAATTAAT 5341  
DB 5141 TTGCTGAAGCGAAACAAACCTGCTGAATTAACAACTCGGATCAAAATGTTAAATC 5200  
QY 5342 ATCAAGATGAAAAATATCAATGTGACTGCAAAAGATTAATTAATGCAATTAATTAAT 5401  
DB 5201 AAGCTACTCTTAAGATGCAATTTGAAGTTCAATTAATTAATTAATTAATTAATTAAG 5260  
QY 5402 CAGTGAAGTTGAGAGCAAAAGAGCGCTGTGCAAGAGAGCTTCTGCAAGTACTACT 5461  
DB 5261 ATTTACCAATTCACACAGGTAATAAGATCAAGTCAACAGATTAATTAATGCTTAATG 5320  
QY 5462 TGAATTAAGACGTTCTTCTCATGTTGATCAACATGATTAATTAACAAAGATTTAAGAG 5521

Db 5331 ATCAGAGAAAATATATATTTTCAGCTGACACTTAATGCAACAGAGATGAAAAAGCAACAG 5380  
Qy 5522 AAAATTAATGAATAAAGAAAAGCAAAATGTTAATGTTCTAGCTGAAAATAGAGTCAAG 5581  
Db 5381 CAATTAAGCAAGTTGACAAAATG--TTCAAACTGCAATTAAAGCATTAATTAATGGTG 5437  
Qy 5582 TGGTCAACAATGCGACAGTCTTTCGAGCAAGTGCACAGCTCAGTAGAGCTGAG 5641  
Db 5438 TGGATTAATGGTACGTTGATGATGATTAACCAAGTTAAAGCAGAAATGATGCTATTC 5497  
Qy 5642 TAGCAGTTAATAAATTACAAAATACTTTCGACATATTAATAAATGACTCAAAATG 5701  
Db 5498 AAGTAGTGTCTACTGTTAAACCTTAAAGCAACCAAGCTATTGAAGTTAAAGCAGAAATG 5557  
Qy 5702 TAGCAATGCTTGGTAAAG-----CAATCTCATTCATCA 5740  
Db 5558 CGAAAGATCTATTTGATCAAAAGTGACCAAGTTAACTGTAAGAAAAAACTGAAGACTTAG 5617  
Qy 5741 TTAAACAATTGGAATTGAGCTGAGTTGAGCTGAGAGAGCTGAGAGCTGAGAGCTGAG 5800  
Db 5618 CAATGATTAAACAATTACAGATCAAGTAAACAAGTATTACTGATGCAACAAGACTG 5677  
Qy 5801 TAGCAGTGAATAGATTGTAATATAGCATAGCAGAAATTAATCATGCAAAATCAGCTG 5860  
Db 5678 CTGAAGTTGAAAAGGAAAGCTCAAGAGCTGAAGCAATTGATTAACATTCAAATCGACT 5737  
Qy 5861 CGAAGGAAATGTCGAGATTATTAACAGAGTCTGATGCGGTAATGCTAATATGAGAGAA 5920  
Db 5738 CAACGAAAAACAAAAGCTTATCGAAGATTAGAAATCTGACTAGACAGATTGAAGGAG 5797  
Qy 5921 CAGTCTGAGATGCGCCGTGAGCAATAGAGAGCTCAACAGTGTATGAAATTAACG 5980  
Db 5798 GTGTAATGTCAACGCTGATGCTAACAACAGAAAGAAAGAGCTTTACGATGCTTAG 5857  
Qy 5981 GATCTCAAAAAGCATATGTAAGATTCTACAGTATGCTTAAAGAGAAACAGATGATT 6040  
Db 5858 AAGACATTTTATCAAAAAGCACTGAAGATTTCTGATCAAACTCAATGACAGAAATG 5917  
Qy 6041 AATTAATCTCTCAAGGCGCAAGTAGAATAAGTGTATAGTAAGTATTCAAAAATCTTAATA 6100  
Db 5918 CTACTGTCAAAAATAGTGGCTTGAACAATTAAAGCACACGTAATTAATCTGAAGTTA 5977  
Qy 6101 TTACGAGAACTTATCACAAAAAGAAAATAGTATTAATAAAGATTTGTTACCAATA 6160  
Db 5978 AAAAAAATGCTTTGGAAGCAATCAGAGAGTGTATACAGCAAAATAGAAATTAATTAATA 6037  
Qy 6161 GTTCAGCTACTCACTT 6178  
Db 6038 ATGCAGATGCAGATGCAT 6055

RESULT 20  
US-10-470-048B-201  
; Sequence 201, Application US/10470048B  
; Publication No. US2005003744A1  
; GENERAL INFORMATION:  
; APPLICANT: MEINKE ET AL.  
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF  
; TITLE OF INVENTION: ANTIGENS TO A SPECIFIC PATHOGEN  
; FILE REFERENCE: SONN:035US  
; CURRENT APPLICATION NUMBER: US/10/470,048B  
; CURRENT FILING DATE: 2003-07-25  
; NUMBER OF SEQ ID NOS: 603  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 201  
; LENGTH: 7434  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
US-10-470-048B-201

Query Match 0.8%; Score 79.2; DB 8; Length 7434;  
Best Local Similarity 41.3%; Pred. No. 0.0029;

Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;  
Qy 3782 TTCAAGTAAACAGAGCTTGGTTAAATAATAGTATTCGAATATCCAAATAATTTA 3841  
Db 3659 TTAATTAATAAAGCAAGCTGACGTATGAAATCAAGCAATTTTAATTAACAAATTTGCAAG 3718  
Qy 3842 AAGTAATGCTTTGAGTGAAGAACTCAAGTAGCAGAGAGAGAGTTGGAAGCACTTA 3901  
Db 3719 AGATTCAAGCTAACGCGATGCAACAGATGAAGAAAAAAGAGAGCTGATGCTGAAGCA 3778  
Qy 3902 AAGAAAGTGAAGAGCAAGAAAGAAAGTTATCTATTGGGAACCTTCCTCTTCAACTTAG 3961  
Db 3779 ATACTGAAATGTGAAGCAATCAAGCATTTCAGAGCAACTTACACCAACAGTTG 3838  
Qy 3962 TGAACAATGAGTTTCTGCAAAATCAGAAATTAATACAGTAGCAGAGAAATCTGAAGACC 4021  
Db 3839 ATGAAGCTTAAGCAAAATGCAAGAGAGGATTTAATGCGGTAAACCAAAAGTTGGAAG 3898  
Qy 4022 AAAAAATGATGTTGATGTCATGCTTTATCAAGCGGACACCAAGTGAACAGAGCTTTAA 4081  
Db 3899 AACAAGCGGCTAAGAGATGAATGATCAATTACAAAGCAACGCAACAAATGTTATCAATA 3958  
Qy 4082 ATTTACAAAGCTGGAAGTCAATGGAATGGAAGGAGCTACTGTGACTGTTGCCAAATTA 4141  
Db 3959 ATGATCAGAACGCTACCAACAGAAAGAAAGAGAGCTATTCAACATTTAGCAACAGAG 4018  
Qy 4142 ACAACAAGTAAATGCTCTTATTAAGTGGTGAAGATTAATCAAGTTAATGAGCGGAG 4201  
Db 4019 TTACAGAGCGGAAATTAATTTACAGCTGCACTGATGATATGTTGTATGAGCGGA 4078  
Qy 4202 CAAGAAGCTTTTGAACAACCTCAAGTACTGCTGAGTACGAGCGGAGGAGCAATTA 4261  
Db 4079 AAGAGCTGGAAGAAATTCATTCAAAGCAAGCAACGCAACAGCGGTTAATCAATG 4138  
Qy 4262 GTTCTGAGCGGAGATTAGAAATTTTCAAGGCGCTTTCTGTCAATTAAGTTGCAATG 4321  
Db 4139 CTAAAAATGATGTGATCACTGACACTGACAACTCAAAATCAAGCAATTAATATACACTG 4198  
Qy 4322 ACGTGAAGCTAGGCTGATTAATCTTCATCGAAGAGAGTAAATGCAATGATCATTTG 4381  
Db 4199 GTGCTACAACTGGAAGAGAAAAATGCAAGAAAGTTTAGTTTAAGCTTAAGAAAAAG 4258  
Qy 4382 CCAAGATGTCAAAGAAAGTTCTGATCTAGCAAAAAGAAATCAAGCTTTACTAAATGAA 4441  
Db 4259 CGTATCAAGATATCTTAATGACAAACAACTAATGATGTTAGCGAAATTAAGATCAAG 4318  
Qy 4442 AAGATTAATAATTTTGAAGATGCTGTATTAATACAGCTGGAATGTTATTAACGA 4501  
Db 4319 CAGTGTGATATTCAGAGTATTACTGCAAGTACAACTAATAAGATGTTG-----CGA 4372  
Qy 4502 AAGAACACTAGAAAAAGCAAAAGAAAAAGAAAGAGCGGTCATTGTAATGCTGCTTAT 4561  
Db 4373 AAGATGAATTTAGCAACAAAGCAAAAGCAAAAGAGCGCTATTGCAAACTGCAAGATG 4432  
Qy 4562 CGGTGCTGGAACGATTAATCCGCTGAGAGTAGTATTGCACTAATCTGTTAAA 4621  
Db 4433 CGACTACTGAAGAAAAAGAAACAAGCAATCAACAAAGTAGAGCAATTAACACAAGGTA 4492  
Qy 4622 AATAATTAAGCAGAAATTTGAGTGAAGCATTAAGAAAGCGGAGAGAGATTAATTCATG 4681  
Db 4493 ATCAAAATATTGAATAATGCAAGTCAATCGATGATGAATGAACATGCAAAAGTAAATGCA 4552  
Qy 4682 CGAAACATGTAATGATGAGGCAAAATCATGCTACTGTTGTTGTAATGCGGCTTGAGC 4741  
Db 4553 TTCAAGCAATTGACCCCAATTCAAGCATCAAGATGTTTAAACGAATGCAAGAGC-GGAA 4611  
Qy 4742 TTGCTATCAGCAAGAGTCTTTTTCAGAGATGGAATCTGAGCATGAGCAAGCTTATCA 4801  
Db 4612 TTGCTAATGAAATGCAAAATTAATACTGAATTAATTAATTAATGAGAGCT----- 4665  
Qy 4802 ATGACCAATGCAAGAGTGTATTAAGAGAAATTTGCTGATATCTTAAATGTAACG 4861  
Db 4666 -----ACTAATGAAGAAAAAGTAAAGATATTGAGCAGTTAGAGCAGCATATGAAGAG 4720

QY 4862 CAATTAATTCATCTTGGGGGTAATGTTGGGGAAACATTCGCGGTTCTCTTTCTACGG 4921  
DB 4721 GTTTAAATTAATTAATGAGCACTACTACAGGTGATTAATCTACTGTAAAGATACG 4780  
QY 4922 CGGTGAGAGCTCTTTGGGAATTAATCTCTCATTAATTAACCTCTGCTTTGATTACG 4981  
DB 4781 CAGTCAAAAAGTTCAACAATCTTCAATCTCTTTAGAAACAGCAGAGTTAAAAAG 4840  
QY 4982 GAACGAGTAATCTTTTACTGGAAGAAATACAAAAGTCAATGTACAGCTTTGAATG 5041  
DB 4841 AATTGATCAAGCTCAGCTGATTAAGAAAACAATAGAACAAACACCAATGATCAGC 4900  
QY 5042 ATTCTCATTTTCAAAACGTTTCTGCTGGAGGCGCTGCAAGTATTAGCAGCTGGAATG 5101  
DB 4901 AACAAATTTAATGATGCAAAAACAAGGTTGATCTGAATTAATTAATCAAGGAAAACA 4960  
QY 5102 GAGGAATGGTATCTGTCAATCGGTTCGTATGAAAACGAGCTTTAGTTGATGATTCG 5161  
DB 4961 ATGTGATCAATCAATCAAAATGAAATATGTTGATTAATGATTAAGAGAAAAGCTA 5020  
QY 5162 AGTTGAAGAGTAATCTTTCAATGTATGCAAAAGATCAAAAACAAATMAATACAA 5221  
DB 5021 AATTATATGCAATTAACATTTAGTGAATACAAAAGATGCTTTAGCTAAATTTGAAG 5080  
QY 5222 TTGCGGAAATGCAAAATGAGAAAACGGCTGAGCTGAGCAAGTCTCATACAA 5281  
DB 5081 ATGCATATATATCTAAAGTAAACGAGCGAATACCTTAACCATCTGACTTCAAGTAAA 5140  
QY 5282 AATTGGAAGAAACAATGATTAGCAATTTGTAAAAAACGTAATTAATCAACGGGGAATG 5341  
DB 5141 TTGCTGAACGAAACAAAACCTTGCTGAATTTAAACAACTGGGATCAAAAATGTTAAATC 5200  
QY 5342 ATCAAGATGAAAAATATCAATGTGACTGCAAAAGATTAATCTATGACCAATACATATAG 5401  
DB 5201 AAGCTACTCTTAAGATGACATTTGAAGTTCAATTCATATGATCTAGATTAATATTAAG 5260  
QY 5402 CAGTGGAGTTGGAGAGCAAAAAGAGCCTCTGTGCAAGAGCTTTCGAAGTACTACCT 5461  
DB 5261 ATTACCAATTCACACAGGTAAAGAAATCAGCTAACACAGTTTATATGCTTAAGCAG 5320  
QY 5462 TGAATTAAGCAGTTTCTTCATGTTGATCAACATGATTTGCAAAAGTTTAGAGAGAG 5521  
DB 5321 ATCAGAAAGAAATTAATTTTCACTGACACTAATCAACACAGATGAAAAAGCAACAG 5380  
QY 5522 AAAATTAATGGAATAAGAAAAAGCAATGTTATGTTCTAGCTGAAAAATACAGATCAAG 5581  
DB 5381 CAATTAAGCAAGTTCACAAAAG---TTCAAATCTCATTTAGAAAACATTTAATATGATG 5437  
QY 5582 TGGTCAAAATGCGACAGTCTTTCGAGAGCAAGTGAACAGCTGAGTGAAGCTGAG 5641  
DB 5438 TGGATTAATGATGCGTTGATGATGATTAACAACAGGTAAAGCAGCAATTAATGATGCTAATTC 5497  
QY 5642 TAGCAGTTAATTAATTAACAAAATTAATCTTTCGACATTAATTAATTAATTAATTAATG 5701  
DB 5498 AAGTAAATCTACTGTTAAACCTTAAGCAACCAAGCTAATGAAAGTTAAAGCAAGAGATA 5557  
QY 5702 TAGCAATGCTTGTAAAAAG-----CAATCTCATTCATCTTA 5740  
DB 5558 CGAAAGAAATCTATGATCAAAAGTACCAAGTTAACTGCTGAAGAAAACATGAAAGCATTTAG 5617  
QY 5741 TTAATAACAATTGGAATGAGCTGAGTTGAGAGCTGAGAGAGCTGAGTGAACAGGTTCTG 5800  
DB 5618 CATGATTAATAACAATTAACAGATCAAGCTTAACAGGTATTAATCTGATGCAACCAACATG 5677  
QY 5801 TAGCAGTAATTAATGTTAATTAATTAATGATAGCAATTAATTAATTAATTAATTAATCAATC 5860  
DB 5678 CTAAAGTTGAAAAAGCAAGCTCAGAGACTTAAGGATTTGAATTAATTAATTAATTAATTAATG 5737  
QY 5861 CGAAGGAAATGTCGAGGTTATTAACAGAGCTGATGCGGTAAATTTGCTAATTAATGACAGAA 5920  
DB 5738 CAACAGAAAAACAAAAAGCTATCGAAGAAATTAAGAACTGCACTAAGACCAATTAAGAACAG 5797

QY 5921 CAGTGTCTGAGTGGCCCGTCAGCAATAGAGCTCAACCAAGTGTGATGAATTAATACAG 5980  
DB 5798 GTGTAATATGTCAGCTGATGCTTACACTGAAGAAAAAGAGCGTTTACGATGCTTAA 5857  
QY 5981 GATCTACAAAAGCATATGTAATAAGTTCTTACAGTATGCTTAAGAGAAACAGATGATT 6040  
DB 5858 AAGACATTTTATCAAAAAGCACTGAAAGATATTTCTGATCAATCAATGACAGAAATCG 5917  
QY 6041 ATATTACTACTCAAGCGCAATAGATTAAGTATGAATTAATTAATTAATTAATTAATTAATTA 6100  
DB 5918 CTACTGTCAAAAATTAATGCGCTTGAACACTTAAGCAACAGTATTAATCTGAAGTTA 5977  
QY 6101 TTAACGAGACTTATCACAAAAAGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6160  
DB 5978 AGAAAAATGCTTTGGAACCAATCAGAGAGTGTAAACAGCAAAATTAATTAATTAATTAATTA 6037  
QY 6161 GTTCAGTACTCATTAATCTT 6178  
DB 6038 ATGCAATGATCAGATGATCAT 6055

## RESULT 21

US-09-815-242-8869  
; Sequence 8869, Application US/09815242  
; Patent No. US2002061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haseibeck, Robert  
; APPLICANT: Ohlsen, Karl U.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8869  
; LENGTH: 7437  
; TYPE: DNA  
; ORGANISM: Staphylococcus aureus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(7437)  
US-09-815-242-8869  
Query Match 0.8%; Score 79.2; DB 3; Length 7437;  
Best Local Similarity 41.3%; Pred. No. 0.0029;  
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;  
QY 3782 TTCAAGTAATAACAAGTCTTTGTTGTTAAATAATGATATTTGGAATGCAATTAATTTTA 3841  
DB 3659 TTAATTAATAACAACAACTGACGTAAATGAATCAACAGCAATTTTAAATTAACAAATTTGCAAG 3718  
QY 3842 AAGTAATGCTTTGAGTGAAGAACTCAAGTAGACAGACAGAGGTTTGAAGACAGTTA 3901

Db 3719 AGATTCAAGCTACGCGCAGATGCAACAGATGAAGAAAAACAAGCACTGATGCTGAAGCAA 3778  
Qy 3902 AAGAAAGTGAGAGACAAGAAAAAGTTATCTATTGGGAACCTTCGCTTCTATCAACTAG 3961  
Db 3779 ATACTGAAAAATGTTAAAGCAAAATCAAGCCATTTCAGCGCAACTCTAAACGACAAGTTG 3838  
Qy 3962 TGAACAAATGAAGTTCTGCAAAATCAGAAATATATACAGTAGCAGAGATCTGAAGCC 4021  
Db 3839 ATGAAGCTTAAGCAAAATGCAAGAGCGATTATGCGGTAAACCAAAAGTTGTAAGA 3898  
Qy 4022 AAAAAATGATGTTGATGTCACCTGCTTATCAAGCGGACACCCAAAGTACAGAGCTTTAA 4081  
Db 3899 AACAAACCGCTAAACATGAATTTGATTCATTTCAAGCAACGCAAACAATGTTATCAATA 3958  
Qy 4082 ATTCAAGCTGGAAGTCAAAATGGAACGTGAGGGCTTACTGTCATGTTGCCAATTAA 4141  
Db 3959 ATGATCGAAGCCGTACACAGAAAGAAAAAGACAGCTATTCAACAAATTAGCAACAGAG 4018  
Qy 4142 ACAACAAAGTAATCTCTTATTAAGTGTGAGATATCTAACGTTAATCGAGCGGACG 4201  
Db 4019 TTACAGAGCGCAAAAATATATATTACAGCTGCAACTGATGATATAGTGTAGTCAAGCGA 4078  
Qy 4202 CAAAAGCTCTTTAGCAACCACTCAAGTGACTGCTGCAAGTGAAGAAGGAGGACAAATTA 4261  
Db 4079 AAGACGCTGGAAGAAATTCATTAATCAAGCACCAACGCAACAGCGGTTAATCAAAATG 4138  
Qy 4262 GTTCTGAGCGGGATTAGGAATTTATCAAGGGGCTGTTCTGTCATTAAGATTGACAAATG 4321  
Db 4139 CTAATAATGATTTGATCAAGCTGAGACACTCAAAATTCACAAATTTGATTAATACACTG 4198  
Qy 4322 ACGTGAAGCTACGCTTGATTAATCTTCATCGAAGAGACTAATGAATCAATGTCATTG 4381  
Db 4199 GTGCTACAACTGAAGAAAAATGCAAGCAAAAGATTAGTTTAAAGTAAAGAAAAAG 4258  
Qy 4382 CCAAGATGTCAAAGAGTTTGATCTAGCAAAAAGATATACAGCTTTACTAAATGGA 4441  
Db 4259 CGTACCAAGATATCTTAATGACAAACAACTAAGATGTTACGCAAAATTAAGATCAAG 4318  
Qy 4442 AAGATAAAAAATATTAGAAAGATCGTGTATTAATACGCTGGAATGTTATTATAGA 4501  
Db 4319 CAGTTGCTGATATTCAAGGTATTACTGCAAGATACAACTAATAAGATGTG-----CGA 4372  
Qy 4502 AAGAACAACTAGAAAAAGCAAAAGAAAAAGAGAGCGGTCATTGTAATGCTGCTTAT 4561  
Db 4373 AAGATGAATTGCAACAAAGCAAGCAAAAGCAAAAGCGTTATTGCACAACTGCAAGT 4432  
Qy 4562 CGGTTGCTGGAACGGAATAATCGCTGAGAGAGTACTATTGCAATCAATCTGTTAAA 4621  
Db 4433 CGACTACTGGAAGAAAAAGCAAGCAAAATCAACAAGTAGACGCAATTTAACACAAGTA 4492  
Qy 4622 ATTAATTTAAAGCAGAAATTGAGTGAAGCAATTAAGGAAGCGGAGAGAGATTAATTCATG 4681  
Db 4493 ATCAAAATATTGAAATGCAAGTCAATGATGATGTAACACTCAAAAGTAATGCA 4552  
Qy 4682 CGAAAATGTAATGTGAGGCAAAATCATCTACTGTTGTTGTGATGCGGCTTGTGAC 4741  
Db 4553 TTCAAGCAATTGACCAATTCAGAGATCAACAGATGTTAAACGAATGCAAGAGC-GGAA 4611  
Qy 4742 TTGCTATCAGCAAAATGCTTTTTCAGGAATGGATCTGAGCATGCGCAGACTTATCAA 4801  
Db 4612 TTGCTACTGAAATCCAAATTAATAATTAATGAAATGAGACT----- 4665  
Qy 4802 ATGAACAGCTTGCAAAAGTGATTAAGGAAGAAATTCGCTGATTCCTTAATGTGAACG 4861  
Db 4666 -----ACTAATGAAGAAAAAGGTAACGATTTGGAACAGTTAGACAGCATATGAAGAG 4720  
Qy 4862 CAATTAATTCATCTTGTGGGTGAATGTTGCGGGAACATTCGCGGTTCTTTTACAG 4921  
Db 4721 GTTAAATATATATTTATGCAAGCACTACTACAGGTGATGTAATCTACTGTAAGATACAG 4780  
Qy 4922 CGGTGAGAGCTGCTTTGCGAATTAATCTTCAATTAATAAACTCTGCTTTGATTACAG 4981

Db 4781 CAGTACAAAAAGTTCAACAACTTATGCAATTCCTGTTAAGAAACAGACAGTAAAAAAAG 4840  
Qy 4982 GAACGAAGGTAAATCCTTTTATGTAAGAAAGATTAACAAAAGTCAATGTATCAAGCTTGATG 5041  
Db 4841 AATTAGATCAAGCTGACAGCTATTAAGAAACAAATTAAGAACAAACACCAATGCATAC 4900  
Qy 5042 ATTCATATTTACAACAGCTTTCGCTGAGGCGCTGCAAGTATTAAAGCGCTGGAATCG 5101  
Db 4901 AACCAAGAAATTAATGATGCAAAAACAAAGAAATGTAATCTGAATTAATCAAGCAAAAAACA 4960  
Qy 5102 GAGGAATGATCTGTCAATGCTGTTGATGTAAGACGAAGCTTTAGTATGATTTG 5161  
Db 4961 ATGTGATCAATTCATCAACAAATGAAATGTTGATTAAGCAGTTAAAGAAAAAGCTA 5020  
Qy 5162 AGTTGAAGAGTAAAGTTCTTCAATGTATGCAAAAGATCAAAAACTAATTAATCA 5221  
Db 5021 AAATTAATGACGTTAAACATTTAGTGTACAAAAGAAAGCTTTAGCTTAATTAAG 5080  
Qy 5222 TTGCGGAAATGCAAAATGGAAGAAAAAGCGCTGAGATTGAGCAACAGTTGCTCATCA 5281  
Db 5081 ATGCATATTAATGCTTAAGTAAACGAAGCGGATTAATCTTAAGCATCGACTTCAAGTGA 5140  
Qy 5282 ATATTGAAAAAATCACTGATTATGCTATTGTAATAAAACAGTAAATTAACAAGCGGAATG 5341  
Db 5141 TTGCTGAAGCGAAACAAAACTTGTGATTTAAACAACTGCGGATCAAAATGTTAATC 5200  
Qy 5342 ATCAAGATAGAAAAAATATCAATGTGACTGCAAAAAGATTACTATGACAAATCTATAG 5401  
Db 5201 AAGCTACTCTTAAGATGACATTGAAGTTCAAAATTCATTAATGACTAGATATATTAACG 5260  
Qy 5402 CAGTCGAGTTGAGAGCAAAAGAGCGCTGTGCAAGAGACTTTCGAATCACTACT 5461  
Db 5261 ATTACAAATTCCAACAGGTAAAAAAAGATCAAGTACAAAGATTTATATGCTTATGCA 5320  
Qy 5462 TGAATTAAGACGTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAA 5521  
Db 5321 ATCAGAAAGAAAAATTAATTTACAGCTGACATTAATGCAACAAAGTAAAAAGCAACAG 5380  
Qy 5522 AAAATTAATGAATTAAGAAAAAGCAAAATGTTAATGTTCTAGTGAATAATGCAATCAAG 5581  
Db 5381 CAATTAAGCAAGTTGACCAAAATG---TTCAACTGCAATTGAAAGCACTTAATATGTTG 5437  
Qy 5582 TGGTCACAATGCGACAGTGTCTTCCGAGCAAGTGAACAGCTGCACTAGAGCTGAG 5641  
Db 5438 TGAATTAATGTAAGCTGATGATGATTAACAAAGTAAGAGCAAGCATTTATGCTATTC 5497  
Qy 5642 TAGCATTAATTAATTAACAAAAATCTTGCAATATATAAAATATGTACTCAAAATG 5701  
Db 5498 AAGTGAATGCTACTGTTAAACCTTAAGCAACCAAGCTATTGAAGTTAAAGCAGAAATA 5557  
Qy 5702 TAGAAATGCTTGTGTAAG-----CAAACTGATTCATCTA 5740  
Db 5558 CGAAAAGATCTAATGATCAAAAGTGAACAGTTAACTGCTGAAGAAAACTGAAGATCTG 5617  
Qy 5741 TTAACAAATTTGAAATTTGAGAGCTGAGAGTGAAGCTGAGAGCTGAGTGAACAGTTTCTG 5800  
Db 5618 CAATGATTTAAACAATTAACAGATCAAGCTAAACAAGTAATTAAGTACCAACAACCTG 5677  
Qy 5801 TAGCATGATTAAGATTGTAATTAATTAACAGATGCAAAATTAATTAATGCAAAATCACTG 5860  
Db 5678 CTGAAGTTGAAAAAGCGAAAGCTCAGAGACTTGAAGCTTAATTAACATTCGAATCGACT 5737  
Qy 5861 CGAAGGGAATGTCGGAATTTTACAGAGCTGATGCGGTAATTTGCTAATTAATGACAGAA 5920  
Db 5738 CAACAGAAAAACAAAAGGCTATCGAAGATTAAGAACTGCACTGACCAAGTTGAAGAG 5797  
Qy 5921 CAGTGTGAGAGTCCCGTGCAGCAATAGAGGCTCAACCAAGTGTGATGAATGAATTAACG 5980  
Db 5798 GTGTAATGTAAGCGCTGATCTCAACTGAAGAAAAAGAGCGTTTACGAATGCTTTAG 5857  
Qy 5981 GATCTACAAAAGCATATGTAAGAAATTTCTACAGATGTTGCTTAAGAAAGAAACAGATGATT 6040  
Db 5858 AAGACATTTTATCAAAAGCACTGAAGATATTTCTGATCAAACTACAAATGCAAGAAATCG 5917

QY 6041 ATATTACTCTCAGGGCAAGTAGATTAAGTGTGATTAAGTATTCAAAAATTCTTAATA 6100  
DB 5918 CTACTGTCAAAAATAGTGGCTTGGAACACTTAAGCACAAGTATTAATCTGAAGTTA 5977  
QY 6101 TTAAGGAGACTTATCACAAAATAAGTAATAAAGAGATTGTACCAATA 6160  
DB 5978 AGAAAAAGCTTTGAGAGCAATCAGAAAGTGTTAACAAGCAATTAATAATTAATA 6037  
QY 6161 GTTCAGTACTCATCTT 6178  
DB 6038 ATGCAGATCCAGATGCAT 6055

## RESULT 22

US-08-781-986A-63  
Sequence 63, Application US/08781986A  
Publication No. US20030054436A1  
GENERAL INFORMATION:  
APPLICANT: Charles Kunach  
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
NUMBER OF SEQUENCES: 5255  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,986A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benson, Bob  
REGISTRATION NUMBER: 30,446  
REFERENCE/DOCKET NUMBER: PB248PP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8512  
TELEFAX: (301) 309-8504  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8155 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-781-986A-63

Query Match 0.8%; Score 79.2; DB 2; Length 8155;  
Best Local Similarity 41.3%; Pred. No. 0.003;  
Matches 998; Conservative 0; Mismatches 1178; Indels 42; Gaps 5;

QY 3782 TTCAAGTAAACAGTCTTTGGTTAAATAAGTATTTGAAATGCCAATTAATTA 3841  
DB 5651 TTAAATAAAAGCACTGCAGTAATGAATTCACAGCAATTTTAAATACAAATTCAG 5710  
QY 3842 AAGTAATGCTTTGAGTGAAGAACTCAAGTAGCAGAGAGAGAGTTGGAAGCAGTTA 3901  
DB 5711 AGATTCAAGCTACGCCAGATGCAACAGATGAAGAAAAACAAGAGCTGATGCTGAAGCA 5770  
QY 3902 AAGAAAGTGAGCAAGAAAAAGTATCTATTGGGAACCTTGCTTATCAACTTAG 3961  
DB 5771 ATACTGAAATGTTAAAGCAATTCAGCCATTTTCAGAGCAATCACTAAGCAAGAGTTG 5830  
QY 3962 TGAACATGAAGTTTCTGCAAAATCAGAAAAATATATACAGTAGCAGAGAAATCTGAAAGCC 4021

DB 5831 ATGAAGCTAAAGCAAAATGCAAGAGAGGATTTAATGGGTAAACACCAAAAGTTGGAAGA 5890  
QY 4022 AAAAAATGATGTTGATGCTACTGTTATCAAGCGACACCCAAAGTGACAGAGCTTTAA 4081  
DB 5891 AAACAAGCGGTAAAGATGAATTTGATCAATTAACAAGCAACGCAACAAATGTTATCAATA 5950  
QY 4082 ATTTCAGAGCTGGAAGTCAAAATGGAACGTAAAGGGGCTACTGTGACTGTGGCAATTTA 4141  
DB 5951 ATGATCAAGCGTACCAACAGAGAAAAAGAGAGCTATTCACAAATTAAGCAACGAG 6010  
QY 4142 ACAACAAATTAATGCTTTCTAATAGTGGGAGATATTAACCTTAATCGAGCGGAG 4201  
DB 6011 TTACAGAGCGCAAAAATTAATATTACAGCTGCACATGATGATATGTTAGATTCAGGCGA 6070  
QY 4202 CAAAAGCTTTTTCAGCAACACTCAAGTACTGCTCAGTGACGACGCGAGGAGCAATTA 4261  
DB 6071 AAGACGCTGAAAAGAAATTCAAATTCAAACAGCAACAGCAACAGCGTTAAATCAAAATG 6130  
QY 4262 GTTCTGAGCGGATTTAGGAATTTCAAGGGGCTGTTCTGTCAATAAGATTGACAATG 4321  
DB 6131 CTAAAAATGATGTTGATCAAGCTGTGACAACTCAAAATCAACAAATGATTAACAATG 6190  
QY 4322 AGTGAAGCTAGCGCTGTATAAATCTTCATCGAAGAGCTAATGAATCAATGTCATTTG 4381  
DB 6191 GTGCTACACTGAGAGAAAAATGCAAGCAAAAGATTGTTTAAAGCTAAAGAAAAAG 6250  
QY 4382 CCAAAAGTGTCAAGAAAGTTCTGATCTAGCAAAAGAAATATACAGCTTTACTTAATGCA 4441  
DB 6251 CGTATCAAGATATCTTAATATGCAAAACAACTAATGATGTTACGCAAAATTAAGATCAAG 6310  
QY 4442 AAGATAAAAATTTTGAAGATCGTGTATTAATACAGCTGGAATGTTTATTAACA 4501  
DB 6311 CAGTTCGATATTTCAAGGTATTTACTGAGATACAAACATTAAGATGTTG-----CGA 6364  
QY 4502 ACGAACATAGAAAAAGCAAAAGAAAAAGAGAGCGGTCAATTTGAATGCTGTTTAT 4561  
DB 6365 AAGATGAATTTAGCAACAAAGCAAAAGCAAAAGCGGTTATTTGACAAATGCGAGAG 6424  
QY 4562 CGGTGCTGGAACGATTAATCCGTGAGAGTACTTATTCAGTCAATCTGTTTAA 4621  
DB 6425 CGACTACTGAAGAAAAAGAAAGAACGAAATCAAGTAGACCAATTAACACAGATA 6484  
QY 4622 ATTAATTTTAACAGAAATTTGATGGAACATTAAGAACCCGAGAGGATTAATCATG 4681  
DB 6485 ATCAAAATATTAATAATGCAAGTCAATGATGATGTAACACTGCAAAAGATTAATGCA 6544  
QY 4682 CGAAACATGTAATGAGAGCAAAATCATCTACTGTTGTAATGCGGCTTTCGAG 4741  
DB 6545 TTCAAGCAATTAACCAATTCAGCAATCAACAGATGTTTAAACGAATGCAAGAGC-GGA 6603  
QY 4742 TTGCTATCAGAAAGATGCTTTTTCAGAAATGGAATCGAGATCGAGCAATGCAACTTATCA 4801  
DB 6604 TTGCTATCAGAAATGCAAAATTAATAATTAATGAAATCTTAATAATATAGACT----- 6657  
QY 4802 ATGACAGATTGCAAAAGTGAATTAAGGAAGATTTCTGCTGATTCCTTAATGTGAACG 4861  
DB 6658 -----ACTAATAGAAGAAAAAGTTAAGATTTGACCACTGATGAGAGATATGAAGAG 6712  
QY 4862 CAAATTAATTCATCTTGGGGTGAATGTTGCGGGAACCAATGCGGCTTCTTTTCACGG 4921  
DB 6713 GTTTAATAATTAATTAATGACCACTACTACAGTATGTAATTAATGTAAGATACAG 6772  
QY 4922 CGGTAGAGCTGCTTTTTCAGAAATTAATCTTCAATAATAAACCCTGCTTTGATTACAG 4981  
DB 6773 CAGTACAAAAAGTTTCAACAACCTTCATGCAATCTGTGTAAGAAACAGCAGGTAAAAAG 6832  
QY 4982 GAAAGAGTAAATCTTTTATGGAAGAAATTAACAAAGTCAATGATGAAGCTTTGAATG 5041  
DB 6833 AATTAGATCAAGCTGAGCTGATTAAGAAAAACAAATTAAGAAACAAACAAATGATCAC 6892  
QY 5042 ATTCTATTTTCAAGCTTTGCTGAGGCGCTCAAGTATTAAGCAGGCTGATATG 5101

```

Db      6893 AACAGAAATTAATGATGCAAAAACAGAAAGTTGATCTGAAATTAATCAAGCAAAAACAA 6952
Qy      5102 GAGGATGATCTCTGCAATCGTGGTCTGATGAAACGGAAGCTTAACTAGTGAATTCG 5161
Db      6953 ATGTGATCAATCAATCAACAAATGATATGTTGATTAATGACGTTAAAGAGAAAGCTA 7012
Qy      5162 AGTTGAAAGAGTAAGTCTTTCAATGTAGATGCAAAAAGATCAAAAACAAATTAATACAA 5221
Db      7013 AAATTAATGCAAGTTAAACATTTAGTGAATCAAAAAGATCTCTTAGCTAAATTTGAAG 7072
Qy      5222 TTGCGGAAATGCAAAATGAGAAAAGCGCGCTGAGAGTTGAGCAACAGTGTCTATACA 5281
Db      7073 ATGCAATTAATGCTAAAGTAAACGAAGCGGATTAACCTTAAGCCTGACCTTAAGTGA 7132
Qy      5282 ATATTGAAAAACATCACTTATAGCTATGTATTAATAAATTAATTAACACGCGCAATG 5341
Db      7133 TTGCTGAAGCGAAACAAAACCTTGTGATTAATAAATTAATTAATTAATTAATTAATTA 7192
Qy      5342 ATCAAGATAGAAAAATATCAATGTGACTGCAAAAAGATTAATCTATGACCAATTAATAG 5401
Db      7193 AAGCTACTTCTTAAGATGACATTTGAAGTTCAAAATTCATTAATGACTTAATTAATTAACG 7252
Qy      5402 CAGTGGAGTTGAGAGAGAAAGAGAGCTCTGTGCAAGAGAGCTTCTGCAATTAATTAAT 5461
Db      7253 ATTACACAAATTCACACAGGTAAAGAAATGAGTACACACAGATTTAATATGCTTAATG 7312
Qy      5462 TGAATTAAGCAGTTCTTCTCATGTTGATCAACTGATTAATGACAAAGATTTAGAGAGAG 5521
Db      7313 ATCAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7372
Qy      5522 AAAATTAATGAAAAATAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAAAATACAGAGT 5581
Db      7373 CAATTAAGCAAGTTGACCAAAATG---TTCAAACTGCAATTAAGAAAGCATTAATTAATG 7429
Qy      5582 TGGTCAAAATGCGCAGAGCTTCCGAGCAAGTGAAGACAGCTGAGTGAAGAGCTGAGAG 5641
Db      7430 TGGATTAATGAGCGCTTATGATGATTAATTAACCAAGGTAAACGACAAATGTTGCTTAAT 7489
Qy      5642 TAGCAGTTAATTAATTAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5701
Db      7490 AAGTGAATGCTAATGTTAACTTAAGGCAACCAAGCTAATTAAGTAAAGCAAGATA 7549
Qy      5702 TAGCAATGCTTTGTTAAAG-----CAATCTCATTCATCTA 5740
Db      7550 CGAAAGATCTATTGATCAAGTGAACAGTTAATGCTGAAGAAAAATGTAAGCATTAAG 7609
Qy      5741 TTTAAACAAATTTGAGAGCTGAGAGCTGAGAGAGCTGAGAGAGTGAAGAGTTCG 5800
Db      7610 CAATGATTAACCAAAATTAACAGATCAAGCTTAACCAAGGTATTAATCTGATGCAACACAG 7669
Qy      5801 TAGCAGTAATTAAGATTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5860
Db      7670 CTGAAGTTGAAAGAGCAAGAGCTCAAGAGCTTAAGAGCATTTGATTAACATTCAAATG 7729
Qy      5861 CGAAGGAAATGTCGAGATTATTAACAGAGCTGATGCGGTAAATTCATTAATTAATTAATTA 5920
Db      7730 CAACAGAAAAACAAAAGCTATCGAAGAAATTAGAAATGCACTAAGCCAGATTAAGCAG 7789
Qy      5921 CAGTGTCTGAGAGTGGCGCTGAGAGCAATGAGAGCTCAACAGGTGATGAATTAATTAACG 5980
Db      7790 GTGTAAATGTCACCGCTGATGCTACACAGTAAGAAAAAGAACGTTTTCGAAATGCTTTAG 7849
Qy      5981 GATCTACAAAGCATATGTAATTAAGATTCTACAGATGTTGCTTAAGAGAAAGACAGATGATT 6040
Db      7850 AAGACATTTTATCAAAAGCACTGAAAGATATTCTGATCAAACTCAAAATGCAAGAAATCG 7909
Qy      6041 ATATTACTACTCAAGGCAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAATTA 6100
Db      7910 CTACTGTCAAAATATGTCGCTTGAACAACTTAAGCAACAGTATTAATTAATTCGTAAGTTA 7969
Qy      6101 TTAAGAAAGCTTATCACAAAAAAGAAAAATTAAGTATTAATTAATTAATTAATTAATTAATTA 6160
Db      7970 AGAAAAATGCTTTGAAAGCAATCAGAGAAAGTGTTAACCAAGCAAAATTAAGTAATTAATAA 8029

```

```

Qy      6161 GTTCAGTACTCATACTT 6178
Db      8030 ATGCAGATGCAGATGCAT 8047

RESULT 23
US-10-329-624-63
; Sequence 63, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunesh
;              Gil H. Choi
;              Patrick S. Dillon
;              Craig A. Rosen
;              Steven C. Barash
;              Michael R. Fannon
;              Stephen C. Barash
;              Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PDL1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8155 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 63:
US-10-329-624-63

Query Match      0.8%; Score 79.2; DB 7; Length 8155;
Best Local Similarity 41.3%; Pred. No. 0.003;
Matches 998; Conservative 0; Mismatches 1378; Indels 42; Gaps 5;

```

Db	6893	AACAAGAAATTAATGATGCAAAAACAGAAAGTTGATCTGAATTAATATCAGCCAAAAACA	6952
Qy	5102	GAGAAATGGATATGTCATATCGTGGTTCTGATGAAAACGAAGCTTTAGTTAGTATCTG	5161
Db	6953	ATGTCGATCATCATCAACAAATGAATAATGTTGATATATGCAGTTAAAGAAAGAAAGCTA	7012
Qy	5162	AGTTTGAAGAGTAAAGTCTTTTCATAGTATGATGCAAAAAGTTCAAAAAACAATAATGCA	5221
Db	7013	AAATTAATGCAAGTTTAAACAATTTAGTGAATCAAAAAAAGATGCTTAAAGTTAAATGAAG	7072
Qy	5222	TTGCGGAAATGCAAAATGAGGAAAAAGCCGCTGAGAGTTGAGCAACAGTTCATATCA	5281
Db	7073	ATGCAATTAATGCTTAAGTTAAAGAAAGAGCCGATTAATCTTAACGATCGACTTCAAGTAAA	7132
Qy	5342	ATCAAGATAGAAAAAATATCAATGTGACTGCAAAAGATTATCTATGACCAATATCTATAG	5401
Db	7193	AAGCTACTTCTTAAAGATGACATTTGAAGTTCAAAATTCATTAATGACTTAATTAATTAACG	7252
Qy	5402	CAGTCGAGTTGAGAGGCAAAAGAGCCCTGTGCAAGAGAGCTTCGCAAGTACTACT	5461
Db	7253	ATTACACAATTCACAACGGTAAAAAAAATACGTACACACAGATTTATATGCTTATACAG	7312
Qy	5462	TGAATAGACAGTTTCTTCTCATGTTGATCAAACTGATATTGACAAAGATTTAGAGAG	5521
Db	7313	ATCGAAGAAAAAATTAATTTACGTGACACTATATGCAACACAGATGAAAGCAACAG	7372
Qy	5522	AAAAATTAATGAAATTAAGAAAAAGGCAAAATGTTAATGTTCTAGCTGAAAAATACAGTCAAG	5581
Db	7373	CAATTAAGCAAGTTGACCAAAATG---TTCAAACTGATTTAGAAAGCAATTAATATATGATG	7429
Qy	5582	TGTCACAAATTGCACAGTCTTTCCGAGCAATGGAACAAGTCGAGTGGAGCTGAG	5641
Db	7430	TGGAATATGGGACGTTGATGATGATTAACACAAAGTTAAAGCAATTTGATCTATTC	7489
Qy	5642	TAGCAGTTAATTAATAATTAACAACAAATCTTCTGCATATTAATAAATAGTACTCAAAATG	5701
Db	7490	AAGTATGATGCTACTGTTAACTTAACGAAACCAAGCTATTGAAGTTAAAGCAGAAAGTA	7549
Qy	5702	TACGAAATGCTTGGTAAAAAG-----CAAACTCATTCATCTA	5740
Db	7550	CGAAAGATCTATATGATCAAAAGTGAACCACTTAATCTGTGAAGAAAAAATGMAACATTAAG	7609
Qy	5741	TTAAAACAATTTGGAAATGAGCTGGAAGTGGAGCTGAGAGCTGAGATGACAGTTCTG	5800
Db	7610	CAATGATTTAAACAAATTAACAGATCAACATCAAAACAGGATTAATCATGATGCAACCAACTG	7658
Qy	5801	TAGCAGTAATTAAGATTTGTAAATATACGATAGCAGAAATTTAAATCATGCAAAAACTAG	5860
Db	7670	CTGAAGTTGAAAAAGCGAAAGCTCAAGAGCTTGAAGCAATTTGATTAATCTCAAAATCACT	7729
Qy	5861	CGAAGGGAATATGTCGAGTTATTACAGAGTCTGATGCGGTAATTTCTAATTTATGACGAA	5920
Db	7730	CAACGAAAAACAATAAGCTATGCAAGAAATTAGAAACTGCACTAGACCAAGATTGAAGAG	7789
Qy	5921	CAGTGTCTGAGTGCCCGCTGCAACATAGAGCCTCAACAGTGTGAATGAATTAACAG	5988
Db	7790	GTTGTAATATGTCMAACGCTGATGCTCAACTGAGAAAAAAGAACGCTTTACGAATGCTTAAG	7849
Qy	5981	GATCTACAAAAGCAATGTTAAAGCTTCTACAGTATGTTCTTAAGAAAGAAACAGTATGTT	6040
Db	7850	AAGACATTTTATCAAAAGCACTGAGATATTTTCTGATCAAACTACAAATGCAAGAAATCG	7909
Qy	6041	ATATTTACTACTCAAGGCGCAAGTAGATAAAGTGTAGATAAAGTATTCAAAAATCTTAATA	6100
Db	7910	CTACTGTCAAAAATATGATGCGCTTGAACAACTTAAAGCACACAGTATTAATCTGAAGTTA	7965
Qy	6101	TTTAAGGAAGCTTATCACAAAAAGAAAAATTAAGTAATTAATAAAGAAATTTGTTACATA	6166

Db 7970 AGAAAAATGCTTGGAGCAATCAGAAAGTGTAAACAAGCAATAGCAATTAATAA 8029  
Qy 6161 GTTCAGTACTCATCTT 6178  
Db 8030 ATGCAGATGCAGATGCAT 8047

RESULT 24  
us-09-815-242-4580  
; Sequence 4580, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 4580  
; LENGTH: 7104  
; TYPE: DNA  
; ORGANISM: *Staphylococcus aureus*  
us-09-815-242-4580

Query Match 0.8%; Score 74.8; DB 3; Length 7104;  
Best Local Similarity 41.6%; Pred. No. 0.017;  
Matches 992; Conservative 0; Mismatches 1367; Indels 27; Gaps 7;

Qy 7208 TAAATTAATTTTGGACAGCTGTAGAAAGATAGAAAAATTTCTGAAGAAAAAGAACAGAG 7267  
Db 2702 TCAATGCAATTAACCAACAAGCAAAAGTTAAACCTGCAGCTGATACGGAAGTAGAAAAAG 2761  
Qy 7268 TTTTAAAACTTTAGACGAAGTTAAACAAGAACAGATTAATAATGATGCTAGCA 7327  
Db 2762 CATATTAATACAGTAAACAAGAAATTCAAAAATAGCAATCTTCAACTACAGAAAAAAC 2821  
Qy 7338 AAAAAATCTTCAATCAGCAGGTATTTCTACAGAAAGATACCTTGTAAAAAGCGATAGAG 7387  
Db 2822 AAGCTGCATATACAGAAATTAAGATACTAATAAAGCAAGAACAAATCTTGATGCTG 2881  
Qy 7388 GAGATATCTCAGGAGAAAGAAATTAAGCCATTTGTGAAGCTTGTATATTAATGAAAAA 7447  
Db 2882 CAAATTAACAACAGTATGTAAACAAGCTAAAGACAAATGTATCTGCATTAATTAACAG 2941  
Qy 7448 ATGTAGATTTTCAACAGAGACAAGAAATATATCACTTCTAGTGGTTTGGGAAGCTG 7507  
Db 2942 TACAAAGCGGCAACAATAAGAAATCGATGCTAAAGCGAAATTCGCTAAAAAGCAAGTG 3001  
Qy 7508 CAGGCTTGCTTCCGATCAGGAACAGTGGCAGTTACAAATATTAAGAAATTCGCGAG 7567

Db 3002 AACGTAAACT---GCAATTAAGCAATGAATGATTGCACTACTGAAGAACAAACAGCTG 3058  
Qy 7568 TTACTGTGAATAATTTCTTTGTGAAGAGCGTGAATAATGTTAGATCGATATTA 7627  
Db 3059 CAAAAGATTAAGTTGATACAGCACTAGTTACTGCAAGCGCTGATTAATGCTGCGAG 3118  
Qy 7628 CAGAAATGTTGCTTTAACAGCATATCAAGGTCTGTAGAGACATTGGAAATGAGAGCTG 7687  
Db 3119 CAAATACTGATGTGATGATTAATGCAAAAACACTAATTAAGAAAGTACATCGCAGCATTAAC 3178  
Qy 7688 CCTATGCAGAATTAATTTCTAATGGAAGATCAAAATATGATTTAAAAATTTCTAGCTAT 7747  
Db 3179 CTGATGCAAAATGTTAAACCAACAGCGAAAACAAGCAATGCTGATTAAGTACAGCGCAAG 3238  
Qy 7748 TAGAAAAAATATGATGTTA---TTGTAAAAATTAATTCGAAATTAAGACGCGAAGCA 7804  
Db 3239 AACGCAATGATGCTATATACGTTGCAACACAGAAAGAAAAAGCAGCTCGAAAACAC 3298  
Qy 7805 AAGATTAACCGTAGAGACGGTAGCTGCCGAGCCATTATCTCAAAAGCAAAAGATGAA 7864  
Db 3299 AAGTTCAAACTGAAAAAACAACAGCTGATACAGCAATTGATGTGCACTACAAATGCG 3358  
Qy 7865 TGAATTCAGAGTTGAAATTGAGAAAGACTATTTTCATGAAGAAATAGAGTAAGCC 7924  
Db 3359 AAGTTGAAGCGGCTAAAAATGCAGAAATGCTAAATTTGAAGCAATTCAGCCAGCAACAA 3418  
Qy 7925 CTTCCTAAAGGAATG---GAAGAGAAATCAATGTCTAAAGTGGAAAAAGAAAAACAGAGTGA 7981  
Db 3419 CAATTAAGATTAATGCAAAACAAGCAATTTGCTACGAAAGCAATGAAGTAAACAGCA 3478  
Qy 7982 CTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTAGACAGGCGCAGAAATTAATTCGGAAG 8041  
Db 3479 TCGCTCAAAAGCAAGACATTACTGTGTGAAGAAATTGACGCGCAATTCGAAT---GTAG 3535  
Qy 8042 CAAAAGATGCCGGAAGCTCTTAATTTGAAAGTTAGTACAAAATCCGGAAGAAATATTTTC 8101  
Db 3536 ATAAATGCTGTACACAAGCAAAATTAACAATGAAGCTGTAATAGTCAAAATGATGTAG 3595  
Qy 8102 ATGCAGATTAATGTGAATTAAGCAACACATTAATAATGAAGTAAACAGCAGTTTCTAAAG 8161  
Db 3596 ACCAAGCAAAACAACGTGTAGTAAGTATGATTAAGTAAACACCAAGTTATTAATA 3655  
Qy 8162 CAGTAACAGTTCTGTATTTGGAGAGAGTTGGAGTCAACCAAGCAGAAAGCTAGCTGAG 8221  
Db 3656 AAGCAACAGCAGTTAACAAGACCGAAAAATATATTAACGCTGCAACTGATATATAGTG 3715  
Qy 8222 GTAAAACTATGTAGAAAGTTAGAGAAAGAAATTTGTTCAAGAACAAATGATGAATGCA 8281  
Db 3716 TAGATACAGCAAGAGATGCAAGTAAAAATTCGATTCAAAGTACACACACCAACAGCG 3775  
Qy 8282 TTTCTAAAGTAAAGTTTGGATGAAGTAAAGTAACTGCTAAATCTTCTGTAGATCAG 8341  
Db 3776 TTAATTAATGCTTAATAATGATGTGATCAAGCTGTGACAACTCAAAATCAAGCAATTG 3835  
Qy 8342 GAAATGAGAGGAATTTGCCGAGCAGAGAGTGAATTACTTTCAACAGCAAAAGTATACAG 8401  
Db 3836 ATAAATACACAGTGTGCTACAACTGAAGAAAAAATGACGCAAAAGATTTAGTTTAAAG 3895  
Qy 8402 AATCCGTATGTTGTTACGAAGCAAGATTATGAATAATATGATTAACAACAAAAATATA 8461  
Db 3896 CTAAGAAAAAGCGTATCAAGATATCTTAATATGCAACAAACACTAACATGTTACGCAAA 3955  
Qy 8462 TTTCAAGATGAATGCTCTTGCTTTAATGATTAACAAAGAAATGAAGCGAATTAAGATCTT 8521  
Db 3956 TTAAGATCAAGCAGTTGCTATGTTCAAGGATTTTCTGCAAGTACAAATTAAGATG 4015  
Qy 8522 TAGCGTAGCCGGTGTGCATCACAAGAACAAACAAAGCAATTTACGAGATCAACAAAGT 8581  
Db 4016 TTGCGAAGATGATTTAGCAACAAAGCAAGGAACAAAGCACTTATTCACAAACAG 4075  
Qy 8582 TAACTTTCAACACTGTAATGAGAGAAACGTATCTCAACTTGTGCAAAAGCTTTGGCTA 8641  
Db 4076 CAGATGCACTATGAAAGAAAAAGAACAAAGCAAAATCAAC-----AAGTATAGCAGAAAT 4129

QY 8642 AAAATGAAATTTGGAATGTAAGAACTGAGAGCCTTAGTCGAGCGGAAAG 8701  
Db 4130 TAAGCGAAGTAACTCAAAATATGAAAATGCAAGTCAATGATATGTAACATGCA 4189  
QY 8702 CAGCGGTGAAATTAATCAAGAGTACTACAGAGCATGGTTGACAGAAATGGGAA 8761  
Db 4190 AAGATATGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTC 4249  
QY 8762 TTGGATTAATTAAGAAAGTTCAGAGATTAATGATGTAAGTCAACGAGAGC 8821  
Db 4250 CAAGGCGGAAATGCTAATGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4309  
QY 8822 GAACCAAGAGGCTTGTTCGAAAGATGATTTCTGTGAAAATCAATTCAGGCG 8881  
Db 4310 AGACTACATTAAGAGAAAGGTAAGATTAAGTCAAGTCAAGTCAAGTCAAGTCA 4369  
QY 8882 AAACAAATTCATCCATTAAGATTAAGTCAAGTCAAGTCAAGTCAAGTCAAG 8941  
Db 4370 GTTAAATTAATTAATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4429  
QY 8942 ATGCTTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATG 9001  
Db 4430 CAGTCAAAAGGTTCAAGCACTTCAAGCAATCCGTTAAGAGAGCAGAGTAAAGCTG 4489  
QY 9002 GTATTGAAATGTTGATGATTAATGATTAATGATTAATGATTAATGATTAAT 9061  
Db 4490 CTTTGAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4549  
QY 9062 GACATGCTATGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 9121  
Db 4550 AACAAGAAATTAAGCA---TGAAAAACAAGGTTGATGATGATGATGATGATGAT 4606  
QY 9122 TAAATATTTTGAAGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 9181  
Db 4607 CAAATATGCAAACTTCAAGCAAGTGAATATGTTGATTAATGATTAAGAGAGAA 4666  
QY 9182 CCAATGAGATGATTAAGAAATTTGAGCAAGAGTATGATCTTCTCAATTAATCA 9241  
Db 4667 CTAAATTAATGATGATTAAGCAATTTAGTGAATCAAAAAAGATGCTTAAAGATT 4726  
QY 9242 AAAATTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9295  
Db 4727 AAGCTGATATTAATCTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 4786  
QY 9295 TTGATGAGGCTGAGAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 9355  
Db 4787 AAATGCGAAG 4846  
QY 9356 AAATTAATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9415  
Db 4847 ATCAAGCTACTCTTAAGAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 4906  
QY 9416 ATGATATGCGGATATGATTAATTAATTAATTAATTAATTAATTAATTAAT 9475  
Db 4907 ACGATTAACATTAATTAAGTGAATTAAGTGAATTAAGTGAATTAAGTGAAT 4966  
QY 9476 CCGATGCGCAAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 9535  
Db 4967 CAGATCAAG 5026  
QY 9536 TAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 9581  
Db 5027 AAGCAATTAAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5072

RESULT 25  
US-09-815-242-8291

; Sequence 8291, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.

APPLICANT: Zvekind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
TITLE OF INVENTION: Prokaryotes  
FILE REFERENCE: EITRA-011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8291  
LENGTH: 7107  
TYPE: DNA  
ORGANISM: Staphylococcus aureus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(7107)  
US-09-815-242-8291

Query Match 0.8%; Score 74.8; DB 3; Length 7107;

Best Local Similarity 41.6%; Pred. No. 0.017; Indels 27; Gaps 7;  
Matches 992; Conservative 0; Mismatches 1367;

QY 7208 TAAATATTTTGAAGAGAGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7267  
Db 2702 TCAATGCAATTCACAAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7261  
QY 7268 TTTTAAAACTTGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7327  
Db 2762 CATATTAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2821  
QY 7328 AAAAAATCTTCAATGAGAGAGTATTTCAAGAGATTAATTTCAAGAGAGTGAAG 7387  
Db 2822 AAGCTGATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2881  
QY 7388 GAGATGCTGAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7447  
Db 2882 CAAATCAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2941  
QY 7448 ATGATGATTAACAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7507  
Db 2942 TACAAGCGGCAACATTAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3001  
QY 7508 CAGCTCTGCTTCCGATCAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7567  
Db 3002 AAGCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3058  
QY 7568 TTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7627  
Db 3059 CAAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3118  
QY 7628 CAGAGATGCTGCTTAAAGAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7687  
Db 3119 CAAATCAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3178  
QY 7688 CCAATGAGAGATTAATTTCAATGAGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 7747

Db 3179 CTGATGCAAAATGTTAAACCAACAGCGAAACAGCAATGCTGATTAAGTACAAAGCGCAAG 3238  
 Qy 7748 TAGGAAAAAATATGATGTTA---TTGTTAAAGATTAATCGAATTTGAGACGGGAGCA 7804  
 Db 3239 AAACAGCAATGATCTAATAACGGGTGCAACAGAAAGAAAAACAGCTGGAAACAAC 3298  
 Qy 7805 AAGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCAAAAGCAAAAGATGAA 7864  
 Db 3299 AAGTTCAAACTGAAAAAACAACAGCTGATACAGCAATTTGATGTCACATCAAAATGCA 3358  
 Qy 7865 TGAATTCAGAGTTGAAATTTGAAAGAGTATTTCAATGAGAAAAATAGATACTAGCC 7924  
 Db 3359 AAGTTGAAGCGGCTAAAAATGCAAAATTTGCTAAATTTGAAACAAATTCAGCAGCAACA 3418  
 Qy 7925 CTTTAAAGGAATG---GAAGAAATCAATGTCAAAGTGGAAAAAGAAAAACAAGTGA 7981  
 Db 3419 CAACCTAAAGATATGAGAAACAGCAATTTGCTACAAAGCGAATGAACTTAAACAGCA 3478  
 Qy 7982 CTGCTGAATCTCAAGAGAGCTTCTGTAGAGCAGTACAGGCGCAGAAATTTTCCGAAG 8041  
 Db 3479 TCGCTCAACGCAAGACATTTCTGCTGAAGAAATTTGACGGCGCAATGGAAT---GTAG 3535  
 Qy 8042 CAAAAGATGCCGAAAGCTCTTAATTTGAAAGTTAGTACAAATTCGGAAGAGTATTTTC 8101  
 Db 3536 ATAAATGCTGTGACACAGCAAAATTAACATTTGAAGCTGCTAATAGTCAAAATGATGTAG 3595  
 Qy 8102 ATGCAATTAATGTAATGAGAACACATTAATGAAAGTAAACAGAGTTTCTTAAG 8161  
 Db 3596 ACCAAGCGAAACAACTGTTGAAGCTAGTATTGATCAAGTAAACCAACAGCAATTAATAA 3655  
 Qy 8162 CAGTAACAGAGTCTGTATTTGGAGAGAGTTGAGTCAACCAAGCAGAAAGTACTGCTGAG 8221  
 Db 3656 AAGCAACAGCAATTCAGACCGGAAAAATTAATTAACGCTCACTGATGTAATGTGT 3715  
 Qy 8222 GTAAATCTATGTTAGAGTTGAGAAAGAAATTTGTTCAACAAATGCAATGAAATGCA 8281  
 Db 3716 TAGATACAGCGAAAGATGCAAGTTAAATTCGATTCAAAGTACACACAGCAACAGCG 3775  
 Qy 8282 TTTCTAAAGTAGAAGTTTGGATGAAAGATTAAGTAACTGCTTAATCTTCTGATGATCAG 8341  
 Db 3776 TTTAAATCAAAATCTAAAAATGATGTTGATCAAGCTGTGCAACTCAAAATCAAGCAATTTG 3835  
 Qy 8342 GAATGAGAGAAATTTGCCGAGCAGAGTGAATCTTCTACAGCAACAAATTAATCTG 8401  
 Db 3836 ATAAATCAACTGTGTCTAACAATGAAAGAAAAATTCAGCAAAAGATTTAGTTTAAAG 3895  
 Qy 8402 AATCGTATGTTGTTAGAAAGCAAGATTATGAAATATGATTAACAAAAAATATA 8461  
 Db 3896 CTAAAGAAAAAGCGTATCAAGATATCTTAAATGCAAAACAACTAACATGTTACGCAAA 3955  
 Qy 8462 TTTCAAGATCAATGCTTGTCTTTAAATGATACAAAGAAATGAGGAATATAGATCTT 8521  
 Db 3956 TTTAAAGATCAAGCAAGTGTCTGATGTTCAAGGTTATTCGCAATCAACAAATTTAAAGATG 4015  
 Qy 8522 TAGCGGTAGCCGCTGTGATGCAAGCAAGAAACAAAGCAATTTACGAGATCAAAAGT 8581  
 Db 4016 TTGCGAAAGATGAATTTAGCAACAAAGCAAGGAAACAAACCTTATTTGCAAACTG 4075  
 Qy 8582 TAACCTTACAACTGTAAATGAGAAACGTAATCTCAACTTGTGCAAAAGCTTTGCTTA 8641  
 Db 4076 CAGATGCGACTACTGAAAGAAAAAGAAACAGCAATCAAC-----AAGTATGTCGAAT 4129  
 Qy 8642 AAAAAGAAATTAATGAAATGTTAAAGAACTGAGAGACCTTATGCGAGCGGAAACAG 8701  
 Db 4130 TAACCGAAGTATCAAAATATTTGAAATGACAGTCAATCATATGTTAAACACTGCA 4189  
 Qy 8702 CAGCGTTGAAATTTACAAAGAGTACTACAGAGCAATTTGTTGCGAGAAATTTGGGAA 8761  
 Db 4199 AAGATTAATGCAATTTCAAGCAATTTGACCAATTTCAAGATCAACAGATGTTAAAGAAATG 4249  
 Qy 8762 TTGAGATTAATTAAGAAAGATTTGCAAGATTAATGCAATTTGAAAGTCAACGAGAGC 8821

Db 4250 CAAGAGCGAATTCCTAATGAAATGCAAAATTAATTAATTAATTAATTAATTAATG 4309  
 Qy 8822 GAACCAAGAGGCTTGTGGAAGAAATGATTTTCTGTGAATAATTAATTAATTAATTAATG 8881  
 Db 4310 AGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 4369  
 Qy 8882 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 8941  
 Db 4370 GTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 4429  
 Qy 8942 ATGCTTGAATGAATCTTATGATGATCTACAGGAAAAAGTGTGCTATGTTGGAATG 9001  
 Db 4430 CAGTACAAAAAGTTCAACACTTATGATGCAATCTGTTTGAAGCGCAGAGTAAATCTG 4489  
 Qy 9002 GTATTTGAAATGTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 9061  
 Db 4490 CTTTATGATCAAGCTGCACTGATTAAGAAACAAATTAATTAATTAATTAATTAATTAATG 4549  
 Qy 9062 GACATGCTATTTGTAATCTGCAAAACAGAAATATCAAGCAATTTTCAAGAGCAAAAG 9121  
 Db 4550 AAACAAGAAATTAAGCA---TGCAAAACAGAAAGTTGATGATTAATTAATTAATTAATTAATG 4606  
 Qy 9122 TAAATATTTCTGGAAGAGAGACCTGCACTGCACTGCAATTAATTAATTAATTAATTAATG 9181  
 Db 4607 CAAATATTAACCAATCTTCAACAGATGAATTTGTTGATTAATTAATTAATTAATTAATTAATG 4666  
 Qy 9182 CCAATGAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 9241  
 Db 4667 CTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 4726  
 Qy 9242 AAAATTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 9295  
 Db 4727 AAGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 4786  
 Qy 9296 TTCAATGAGGCTGCTGAAGCAAGAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 9355  
 Db 4787 AAATTTGCAAGAGCAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 4846  
 Qy 9356 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 9415  
 Db 4847 ATCAAGCTACTCTTAAGATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 4906  
 Qy 9416 ATGATATGCGGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 9475  
 Db 4907 ACGATTAACAAATTCGAAGCAAGTAAAGAAATGATGATGATGATGATGATGATGATGATGATG 4966  
 Qy 9476 CGATGCCAAAAGTCACTGCTGAGCTGCTGCGCACTGCACTTAATTAATTAATTAATTAATTAATG 9535  
 Db 4967 CAGATCAGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 5026  
 Qy 9536 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 9581  
 Db 5027 AAGCAATTAACAGATTAACCAAAATTTTCAAACTGCAATTAAGAAAG 5072

RESULT 26  
 US-10-972-079-13288/c  
 ; Sequence 13288, Application US/10972079  
 ; Publication No. US2005015317A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MMI GENOMICS, INC.  
 ; APPLICANT: DENISE, Sue K.  
 ; APPLICANT: ROSENFELD, David  
 ; APPLICANT: KERR, Richard  
 ; APPLICANT: BATES, Stephen  
 ; APPLICANT: HOLM, Tom  
 ; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BREED  
 ; FILE REFERENCE: MM1110-2  
 ; CURRENT APPLICATION NUMBER: US/10/972,079  
 ; PRIOR APPLICATION NUMBER: US 60/514,333  
 ; PRIOR FILING DATE: 2003-10-24











Qy	9072	TGTGAAACCTCTGGAA----	ACAAGAATATCAAGCATTTTACAAAGGCAAAAGCTAATA	9127	
Db	2619	AAAAAAAAAAAAATTAATAATCTTA	AAAAAAAAAAAAACAAAATCTTA	2560	
Qy	9128	TTCTTGGAAAAAGAGACCGCTG	CACTGCAATATCGAATGTACACATTTCCAATG	9187	
Db	2559	TAAAAA	AAAAAATTTTAAAAAATTAATAATCTTA	2500	
Qy	9188	AGATGATATTTAAAAATTTGG	CAAAAGCATGTGATCTTTCATTTAATACAAAAAT	9247	
Db	2439	TAAAAATTAATAATTAATAATTC	AAAAAATTAATAAATCTTA	2440	
Qy	9248	CAAAAAATTAATATTA	CTTTAGCATCAAGTAGATCGAATGTGATGTTTCATGCGGTGG	9307	
Db	2439	TAAAAA	AAAAAATCTTA	2380	
Qy	9308	CTGAAGCAAGAGGTGACGAGAC	CCAAAGCGACATTAAGTAAATCAATTAATAGAA	9367	
Db	2379	AAAAA	AAAAATTTAAAAAATTAATAATCTTA	2320	
Qy	9368	CTAATTAATGTTGATTTAG	CAGAAAAATTTAAACAGAGGAGAAATCATCATGTATATGTCGG	9427	
Db	2319	AAAAA	AAAAATTTAAAAATTAATAATCTTA	2260	
Qy	9428	GATATGATTAATAATTAATTA	TAAAGTAAGCAAAATTTAAGCGTATGCGATGCCAAA	9487	
Db	2259	AAAAA	AAAAATCTTA	2200	
Qy	9488	GTCATGCTGACGCTGCTT	CGGCAACTGCGACATTTGAAAAAATGAAGTAATTT----	9543	
Db	2139	AAAAA	AAAAATTTAAAAATTAATAATCTTA	2140	
Qy	9544	--AATATGCAATCGAGAA	ATTTAAAAATCTGGCAAGTGGAAAGGAAGCTAATA	9601	
Db	2139	AAAAA	AAAAATCTTA	2080	
Qy	9602	AAAAA	AACTCGGTAGATCTAATCAAGTGA	CTGTATACGGATTAATATTCATGGCAT	9661
Db	2079	AAAAA	AAAAATATCTTA	2020	
Qy	9662	CTTCTGA	AAAAAGCATCAAAAAATTTGACATATCAATCAAAAGAGAGCAAAAA	9721	
Db	2019	AATCTA	AAAAAACACTAATTAATAAAGCAATCTTA	1960	
Qy	9722	AA	9723		
Db	1959	AA	1958		

```

; NAME/KEY: CDS
; LOCATION: (2) ..(1891)
; OTHER INFORMATION:
US-09-742-096-4

```

Qy	6864	AGCCGAGATCTGAGAGTGGTTCTGTCAATTAATTGTGCAATTAAGTACAAACAGATAT	6523
Db	958	AGAAAGTGTGCTGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAAGAAAGTGTAGC	1017
Qy	6924	CGATCATAGTACTTTACACTCTTCTACTGATGTAAATGTAAAGCTCTTAATTAATTTTC	6983
Db	1018	TGAAATGTTGAAGAAAGTGTAGCTGAAATGTTGAAGAAATGTGTAGCTCCAGTGTGA	1077
Qy	6984	GAATTCCTTGACAGCCGGTGGAGAGCCGACAGTCTTGACAGCAATTACCGAGTGTTC	7043
Db	1078	AGAAATGTGAGCTCCACTGTTGAAGAAAGTGTGCTGAAAGCGTTCACAAATTTATC	1137
Qy	7044	TGTTAACCTATAATAGTTCTGTGATAGCTCGAGTTCACAATTAATCTGTATTTGACTTC	7103
Db	1138	AGACAACTTTTAAATTAATTTATTAAGTGTGTATGAAACTGTAGGAAATTAAGACAGTAT	1197
Qy	7104	CGTACGAGAAATAATTAATGTAAACGGCAAAAGAGAAATAATTAAGCAACAGAGC	7163
Db	1198	ATTAAATGAGTAAAGAAAGTAAAGAAATGTAGTCCACCAATACTAGAAAGTAGA	1257
Qy	7164	AAATGCAGAAATCCGAGAGCAGCAATCGAGCCAAATGCTTGGTAATTAATTTTGGAC	7223
Db	1258	AGAAACTACAGCTGAAGTGTACTTTTAGTAATATNTAGAGAGATACAAAGAAA	1317
Qy	7224	AGCTGTAGAAATGAAAAAATTCGTGAAGAAAAGAAACAGAGTTTAAAACTTTAGA	7283
Db	1318	TACTATTACTAATGATCTATAGAGAAAAATTTAGAAAGCTCCACGAAATGTATTAG	1377
Qy	7284	CGAAGTTAACAAAGAACAGATTAATAAAGATGCTACGAAAAAATCTTACAAATC	7343
Db	1378	TGCCGCTTTGAAAAATACCCAAAGTGAAGAGAAAAGAAAGTAATTAAGTGAATTGA	1437
Qy	7344	AGCAGGTATTTCTACAGAAAGATCTTCTGTAAAGCGGATGAGAGATCTCAGGAGA	7403
Db	1438	AGAAATAAAGAAAGAGTCGCTTACCACTTTAATGAACCTGTGAACAGGAGAAAGAA	1497
Qy	7404	AGGAATTAAGCCATTGTGAAGACTCTGATATTATTGGAAAAAATGTAGA-----TAT	7457
Db	1498	GAGCGAAAGTACAAATTACGAAATATTGAAAAATTTAGAAAGAAATGACAGTGAAGATTA	1557
Qy	7458	TACAACAGAGCAAGAAATATATCACTTCTACCTGTGTGGTGGAACTGAGAGTCTTC	7517
Db	1558	TGAAAAAGTTCAAGGAATTTAGAAATTTAAAGAAACGTATTTAATACGTATTAGA	1617
Qy	7518	TTCCGCATCAGAAACAGTGCAGTTACAAATATTAAAAAGAAATTCGGAGTTACTGTTGA	7577
Db	1618	TAAAGTAGAGAAACAGTAGAAATTTAGCGGAGAAAGTTTAAAGAAACATGAATGATTA	1677
Qy	7578	AAATCTTTTGTGAAGCAGCTGAAAAAGTAAATGTTAGATCGATATATTACAGAAAAATG	7637
Db	1678	AGCAATTTTGTGTGAATATTGTAAATGTAAAGGAATCAAGAAATTTAATTACAGG	1737
Qy	7638	TGCTTT 7643	
Db	1738	TATGTT 1743	
RESULT 34			
US-10-415-253-3			
Sequence 3, Application US/10415253			
Publication No. US20040067236A1			
GENERAL INFORMATION:			
APPLICANT: Cohen, Joe			
APPLICANT: Druilhe, Pierre			
TITLE OF INVENTION: Immunogenic Compositions Comprising			
FILE REFERENCE: B45250			
CURRENT APPLICATION NUMBER: US/10/415,253			
CURRENT FILING DATE: 2003-04-25			
PRIOR APPLICATION NUMBER: PCT/EP01/12349			
PRIOR FILING DATE: 2001-10-23			
PRIOR APPLICATION NUMBER: EP00203724.0			
PRIOR FILING DATE: 2000-10-25			

Db 989 GTGATCGAACTGAGCAATAAAGACAGATATTAATGAGATAGAGAAAGTAAAG 1048  
Qy 7130 CAAAAGAGAAAAAATTAAGCAACAGCGCAATGCGAGATCGAGAGAGAGCA 7189  
Db 1049 AAAATGTAGTCCCAATCTAGAGAAAAGTGAAGAACTACACTGAAGTGTACTA 1108  
Qy 7190 TCGAGCCATGCTTGGTAAATTAATTTGGACAGCTGTAGAAATGAAAAATTTCTG 7249  
Db 1109 CTTTAGTATATATTAGAGAGATACAAAGAAATATCTATTAATGATCTATAGAG 1168  
Qy 7250 AAGAAAGGAAACAGAGATTAAACCTTAGAGCAAGTTACAAAGAACAGATAAAA 7309  
Db 1169 AAAAATTGAGAACTCCACGAAATGTATTAGCCGCTTAGAAAAATCCCAAGTG 1228  
Qy 7310 AAGTAATGATGCTACGAAAAAATCTTACATCAGCAGTATTTCTACAGAACTACTT 7369  
Db 1229 AAGAGAAAAAGAAAGTAAATGATGTAATTGAAGAAAGTAAAGAGGTCGCTACCA 1288  
Qy 7370 CTGTAAAGCCGATAGAGAGATCTCAGGAGAAAGAAATTAAGCCATTGTGAAGACTT 7429  
Db 1289 CTTTAATAGAACTGTGAAACAGGAGAAAGAGAGCAATTAAGGAAATAT 1348  
Qy 7430 CTGATATTATTGAAAAAATGTAG-----ATATTACAAAGAGAGCAAGAAATTAATCA 7483  
Db 1349 TTGAAAAATTAGAAAGAAATGCAGTAGAAAGTAAAGTTGACAGAAATTTAGAGA 1408  
Qy 7484 CTTCACGTGCTGTTTGGAACTGCAAGCTTCTCCGATCAGAAACAGTGGAGTTA 7543  
Db 1409 AATTAAAGCAACCTGATTTTAACTGATTAAGTAAAGTAAAGAAACAGTGAATTA 1468  
Qy 7544 CAATATTAAAGAAATTCGAGATTACTGTGAAAAATCTTTTGTGAAGCAGCTGAA 7603  
Db 1469 GCGGAGAAAGTTAGAAAAACATGAATGGAATTAAGCATTTTGTGAGAAATTTGATA 1528  
Qy 7604 AAGTAATGTTAGATCGATATTACAGAAATGTTGCTTT 7643  
Db 1529 ATGTAAGGAATACAGAAATTTATTTAACAGTATGTT 1568

## RESULT 35

US-10-425-115-115706/c  
; Sequence 115706, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222) B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 115706  
; LENGTH: 1121  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1121)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_37006C.1  
US-10-425-115-115706

Query Match 0.7%; Score 63.8; DB 8; Length 1121;

Best Local Similarity 44.8%; Pred. No. 0.63;

Matches 290; Conservative 0; Mismatches 353; Indels 5; Gaps 1;  
Qy 7110 AGAAAAAGTAAATGTAAACGCAAAAGAGAAAAATATTAAAGCAACAGCAAAATGC 7169

Db 948 AAAAAAGAAAAATTAATAAAAAATTAATAAAAAAGAAAAAATAAAAAATGAAAA 889  
Qy 7170 AGAATCGAGAGACAGCAATCGAGCCAAATGCTGTGTAATTAATTTGGACGCTGT 7229  
Db 888 AAAAAAGAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAA 829  
Qy 7230 AGAAGTAGAAAAATTCGTAAGAAAAAGCAAGATTTTAAAACTTTAGAGAAAGT 7289  
Db 828 AAAAAAGAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAA 769  
Qy 7290 TAAAGAGACAGATTAATAAGTAAATGATGCTACGAAAAAATCTTACATCAGCAGG 7349  
Db 768 AAAAAAGAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAA 709  
Qy 7350 TATTTCTACAGAGTACTTCTGTAAAGCGGATAGAGAGATCTCAGAGAGAGAAAT 7409  
Db 708 TAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAA 649  
Qy 7410 TAAAGCATTGTGAGACTTCTGATATTATTGAAAAAATGATATTTACACAGAGGA 7469  
Db 648 TAAAGTATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAA 589  
Qy 7470 CAAGATTAATCACTTCTACTGCTGTGGAACAGCAGCTCTGCTTCCGATCAGG 7529  
Db 588 AATTAACATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAA 534  
Qy 7530 AACAGTGCAGTTCAATTAATAAGAAATTCGGAATTCGCTGTGAAAAATTCCTTTGT 7589  
Db 533 AAAAAATTAATAACAAAAACAAAAAATAATAATAATAATAATAATAATAATAATAATAATA 474  
Qy 7590 GAAAGCAGCTGAAAAAGTAAATGTTAGATCGATATTACAGAAATGTTGCTTTAACAGC 7649  
Db 473 AAAAAAGCAAACTATCAAAACCTGATTAACAAATTAATAAGAGTATCTTCAATCC 414  
Qy 7650 ATATCAAGCTCTGTAGAGCATTTGGAAATGAGCTGCTATGCAATTAATTTCTAA 7709  
Db 413 ATATCCATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAAATTAATAAAAA 354  
Qy 7710 TGGAAGATCAATATCAATTAATAAAATTCATAGCTATTAGAAAAA 7757  
Db 353 AACATTAACACACTCAATTAATAATTCCTACATCTGATTAATAAAA 306

## RESULT 36

US-10-425-115-151933/c  
; Sequence 151933, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222) B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 151933  
; LENGTH: 996  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(996)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_70090C.1  
US-10-425-115-151933

Query Match 0.6%; Score 63.2; DB 8; Length 996;

Best Local Similarity 44.0%; Pred. No. 0.76;

Matches 310; Conservative 0; Mismatches 394; Indels 1; Gaps 1;





OY	363	TCGAGGAAATTTATTTCTTAAAGTCGGAGGAGTGGCAGTAGGAAAAATGAGCTTAT	422
Db	39818	TAATAAATATAGACATATATTCAGACATATATTTATTTAAATATTTTAAATATATTTGATTA	39759
OY	423	CAATGCTGCTTCTTTTCATTCTATATTTCCAAAAACAAGATGATTTTAAAGAGCTTTGGA	482
Db	39758	AAATTCATATTCAAAACATATTTATTTAAAAACAGTATTTAAATGTATTAATAATATGTTTTTGA	39699
OY	483	AGAAAGCAAAATGATGTAAGTTTAAATGGAATCATTTCCAGTATGATGAAAAAGTAAAT	542
Db	39698	AAATTAATAATATAGATTAATATATAGTAAAAATGCTTCTATATATATATAGAAAAAATATA	39639
OY	543	TCCATTTGAATCCGATGGAAGCATTTACGGTAGAAGAAAAATCAATGCTGTTGAAGGCAT	602
Db	39638	AAATACATTTTATTTCTCTCATTAATAATAGTTATCAATTGATATGATATTAATAAATCTGG	39579
OY	603	CGGTTTATATGTCGGCGGATATAGATTTGAAGAAGTACTGCATATCTAATAACAGAAATTAC	662
Db	39578	ATTTTTCATGCTTCTGATGTGTGTATTAATAATTTATAGAAAAATATAATATATTTGGCTTAA	39519
OY	663	AGATTTTAAAAATTTAGTCAATATTAATAGTATGCAATTAATTTCTGCTGACCGGAGATT	722
Db	39518	TAATTAATAAAACATATAATAATTAATAATAATTTATATGAATTTATTAATATATTAATAATA	39459
OY	723	AAAAAGTACCAAGCAAAATCTGGAGATTTATTTCTTTCAGCTCACATGAT-----TC	776
Db	39458	TAATTAATATATTAATAATTAATAATAATAGTTGATTAATATGATTAATATTAATGGAATATA	39399
OY	777	TCCTCAAAAAAGCTATGGGAAAAAATTCACGTTTGAAGAAGAAATGAAAGATATGTATA	836
Db	39398	TCCTACATCTATTTTATTTAAATCTTTATTAAGATGATTAATATTTAAATTAATATATA	39339
OY	837	AGGAAATATCCAAAGCAAAATATTTGATCTG-ATGCTGATTTGGAAGCAGATGCAATATATA	895
Db	39338	ATATATATATTTATTTGATTTATATTAATTTAAATGTTTAAATATCAATTTATCTATATATA	39279
OY	896	AAATTTAGTGCMAAGCTACAAATAGGAGATTTTAAAGAAAGAGGGGAAAAAGAACTT	955
Db	39278	AAATATATGAAATTTATTAATAAAAAATTTATTAATTAATAATATGAAAAAAACCTATATATA	39219
OY	956	ATTAACATCTCTTTAAGTTTATCAGATGTGGAAGCTCCGTAAGATTAATTAAGGAAGAAG	1015
Db	39218	ATTTAGCAAAATACATTAATTAATCTATCTATTAATGAAAAAACAAGAAATTAATTTAAAG	39158
OY	1016	TCATGGAAGAAGTGTGACATTAACAGTGAAGCAAGAAATTTCTATGAT	1065
Db	39158	AAGTAAATTAATGATGATGATATAAATGTATATAGAAAAAGATTTTATTAAT	39109

Search completed: November 26, 2005, 06:34:14  
Job time : 4553 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW nucleic - nucleic search, using SW model

Run on: November 25, 2005, 20:31:24 ; Search time 307 Seconds  
(without alignments)  
4708.018 Million cell updates/sec

Title: US-10-647-057-8

Perfect score: 9726

Sequence: 1 atgagcgagcatcaataa.....gagaaaggaataatga 9726

Scoring table: IDENTITY NUC

Searched: 3205263 seqs, 74304013 residues

Total number of hits satisfying chosen parameters: 6410526

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New.\*

1: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.2	0.8	4965	US-10-485-517-43	Sequence 43, Appl
2	63.2	0.6	2407	US-10-485-517-42	Sequence 42, Appl
3	51.4	0.5	3543	US-10-793-626-3493	Sequence 3493, Ap
4	50.2	0.5	2189	US-10-485-517-121	Sequence 121, App
5	50.2	0.5	3174	US-10-485-517-57	Sequence 57, App
6	48.4	0.5	3343	US-10-485-517-12	Sequence 12, Appl
7	47.8	0.5	3543	US-10-793-626-4149	Sequence 4149, Ap
8	47.8	0.5	3840	US-10-793-626-3187	Sequence 3187, Ap
9	45.6	0.5	1293	US-11-074-176-131	Sequence 131, App
10	45.6	0.5	2275	US-10-485-517-11	Sequence 11, Appl
11	45.6	0.5	3309	US-10-793-626-3637	Sequence 3637, Ap
12	45.4	0.5	663	US-10-793-626-1821	Sequence 1821, Ap
13	45.4	0.5	3180	US-10-793-626-3712	Sequence 3712, Ap
14	45.4	0.5	3650	US-10-793-626-3756	Sequence 3756, Ap
15	45.2	0.5	570	US-10-793-626-2229	Sequence 2229, Ap
16	45.2	0.5	2858	US-10-793-626-3777	Sequence 3777, Ap
17	45.2	0.5	3002	US-10-793-626-3968	Sequence 3968, Ap
18	45.2	0.5	3063	US-10-793-626-3432	Sequence 3432, Ap
19	45.2	0.5	3295	US-10-793-626-3676	Sequence 3676, Ap
20	45.2	0.5	3379	US-10-793-626-4138	Sequence 4138, Ap
21	45.2	0.5	3518	US-10-793-626-4114	Sequence 4114, Ap
22	45.2	0.5	3841	US-10-793-626-3497	Sequence 3497, Ap
23	45.2	0.5	3985	US-10-793-626-3856	Sequence 3856, Ap

24	45.2	0.5	4719	US-10-793-626-2905	Sequence 2905, Ap
25	45	0.5	2119	US-10-485-517-13	Sequence 13, Appl
26	45	0.5	2621	US-10-793-626-4448	Sequence 4448, Ap
27	45	0.5	3589	US-10-793-626-3353	Sequence 3353, Ap
28	44.8	0.5	1206	US-10-793-626-835	Sequence 835, App
29	44.8	0.5	3354	US-10-793-626-4188	Sequence 4188, Ap
30	44.8	0.5	3497	US-10-793-626-4025	Sequence 4025, Ap
31	44.6	0.5	3066	US-10-793-626-3818	Sequence 3818, Ap
32	44.6	0.5	3170	US-10-793-626-3849	Sequence 3849, Ap
33	44.6	0.5	4329	US-10-793-626-2051	Sequence 2051, Ap
34	44.4	0.5	1500	US-10-793-626-1557	Sequence 1557, Ap
35	44.4	0.5	3145	US-10-793-626-4198	Sequence 4198, Ap
36	44.4	0.5	3328	US-10-793-626-4195	Sequence 4195, Ap
37	44.4	0.5	3335	US-10-793-626-3503	Sequence 3503, Ap
38	44.4	0.5	3771	US-10-793-626-4066	Sequence 4066, Ap
39	44.2	0.5	2454	US-10-793-626-49	Sequence 49, Appl
40	44.2	0.5	2454	US-10-793-626-1527	Sequence 1527, Ap
41	44	0.5	1044	US-10-793-626-461	Sequence 461, App
42	44	0.5	1044	US-10-793-626-1267	Sequence 1267, Ap
43	44	0.5	3153	US-10-793-626-3348	Sequence 3348, Ap
44	44	0.5	3267	US-10-793-626-4042	Sequence 4042, Ap
45	44	0.5	3618	US-10-793-626-3564	Sequence 3564, Ap

## ALIGNMENTS

RESULT 1	US-10-485-517-43/c	Application US/10485517
Sequence 43, Appl	US/10485517	
Publication No. US20050256299A1		
GENERAL INFORMATION:		
APPLICANT: University of Sheffield		
APPLICANT: Biosynex Incorporated		
APPLICANT: Foster, Simon		
APPLICANT: Mond, James		
TITLE OF INVENTION: Antigenic Polypeptides		
FILE REFERENCE: P100629MO		
CURRENT FILING DATE: 2004-02-02		
PRIOR FILING DATE: 2001-08-02		
PRIOR APPLICATION NUMBER: GB 0118825.9		
PRIOR FILING DATE: 2002-01-09		
PRIOR APPLICATION NUMBER: GB 0200349.9		
NUMBER OF SEQ ID NOS: 424		
SOFTWARE: Patencin version 3.1		
SEQ ID NO 43		
LENGTH: 4965		
TYPE: DNA		
ORGANISM: Staphylococcus aureus		
US-10-485-517-43		
Query Match	0.8%; Score 79.2; DB 1; Length 4965;	
Best Local Similarity	41.3%; Pred. No. 1.8e-05;	
Matches 998; Conservative	0; Mismatches 1378; Indels 42; Gaps 5;	
3782	TTCAAGTAAACAAAGTCTTTGTTAAATAGTATTTGGAATGCCAATAATTAA	3841
4395	TAAATTAACCAACTGCATTAATGAATCAAGCAATTTAAATTAACAATTCCAG	4336
3842	AAGTAAATCTTGAAGTGAAGCACTGAGCAGAGAGGTTTGAAGCAGTTA	3901
4335	AGATTCAACTGCGCAGATGCAACGATGAAGAAACAAAGCAGCTGATCTGAAGCA	4276
3902	AAGAAAGTGAGACAAAGGAAAAAGTTATCTATTGGAACTTCTGCTTATCACTTAA	3961
4275	ATACTGAATGATGTAAGCAATTCAGCACTTTCAGCACTAATGACCAAGTTG	4216
3962	TGAACATGAAGTTTTCGAAATTCGAAATTAATCACTAGCAGAGAAATCGAAAGCC	4021
4215	ATGAACTTAAGCAATGCAAGCAGCAGATTAAATCGGTAAACCAAAAGTTGGAAGA	4156
4022	AAAAATGATGTGTATGTCATGCTTATCAAGCGGACACCAAGTGACAGAGCTTTAA	4081

Db 4155 AACAAAGCGGCTAAAGATGAAATTTGATTCATTTACAAGCAACGCAACCAATGTTATCAATTA 4096  
Qy 4082 ATTTACAAGCTGGAAGTCAAAATGGAATCTGTAGGGGCTACTGTGACTGTGGCCAAATTTAA 4141  
Db 4095 ATGATCAGAACCTTACACACAGAAAGAAAAGAACGCTATTTCACAACATTTAGCAACGCG 4036  
Qy 4142 ACACAAAGTAAATGCTTCTATTAGTGTGTGGAGATATCTAACGTTATCGAGCGGACG 4201  
Db 4035 TTACAGACGCGAAAAATTAATTTACAGCTGCACATCGATGATATAGTGTATGATCAGGGCA 3976  
Qy 4202 CAAAAGCTTTTATGACACACTCAAGTACTGTGCACTGACGACGCGGAGGACAAATTA 4261  
Db 3975 AAGACGCTGGAAAGAAATTCATTTCAAGCAGCAACGACCAACAGCGGTTAATCAATG 3916  
Qy 4262 GTTCTGAGCGGGATTAGGAAATTAATCAAGGGGCTGTTCTGTCAATAGATTGACAAATG 4321  
Db 3915 CTAAAAATGATGTTGATCAAGCTGTGACAACTCAAAATCAGCAATGTGATATACAACTG 3856  
Qy 4322 ACGTGAAGCTAGCGTTGATTAATCTTCATCGAAGAGCTAATGAAATCAATGTCATTG 4381  
Db 3855 GTGCTACACTGAAGAGAAAATGACGCAAAAGATTTAGTTTAAAGCTTAAAGAAAAG 3796  
Qy 4382 CCAAGATGTCAAAGAGTTCTGATCTAGCAAAAAGATATCAGGCTTTACTTAAATGGA 4441  
Db 3795 CGTATCAGATATCTTAAATGCACAAACTAATGATGTACGCAAAATTAAGATCAAG 3736  
Qy 4442 AAGATTAATAATTTAGAAAGATCGGTGATTAATACGCTGGAATGGTTTATTAACA 4501  
Db 3735 CAGTGTGCTGATTTCAAGGATTAATCTGCAGATACAACTAATAAGATGTTG-----CGA 3682  
Qy 4502 AAGAACCACTAGAAAAAGCAAAAGAAAAAGAGAGCGGTCAATGTAATGCTGCTTAT 4561  
Db 3681 AAGATGAATTTAGCAACAAAAGCAACGAAACGCTTATTTAGCAACAACTGCAAGT 3622  
Qy 4562 CGGTGCTGGAACGATTAATCCGCTGAGAGAGTACTTATGCAATCTTAAATGTTAAA 4621  
Db 3621 CCACACTGTAAGAAAAAGAACCAAGCAATCAACAGTAGACCACTAATTAACAAGATA 3562  
Qy 4622 ATTAATTTAAGACAGATTGAGTGAAGCAATTAAGGAAGCGGAGAGATAAATTCATG 4681  
Db 3561 ATCAAAATTTGAAATGACACAGTCAATGATGATTAACACTGCAAAAGATTAATGCA 3502  
Qy 4682 CGAAACATGTAATGTGAGGCAAAATCATCTACTGTTGTTGTAATGCGGCTTGAGC 4741  
Db 3501 TTCAAGCAATTCACCAATTCACAGATCAACAGATGTTAAAGATGCAAGC-GGA 3443  
Qy 4742 TTGCTATCGCAAAAGTCTTTTTCAGGAATGGGATCTGAGCAATGCAAGCTTATCA 4801  
Db 3442 TTGCTAATCTGAATGCAAAATTAATTACTGAATCTTAAATTAATGAGACT----- 3389  
Qy 4802 ATGACACGTTGCAAAAGTGTGAATTAAGAAATTTCTGCTGATTCCTTAAATGTGAAG 4861  
Db 3388 -----ACTAATAGAAAAAGGTAACGATTAATGACAGTATGAGAGCATATGAAGAG 3334  
Qy 4862 CAAATTAATTCATTTCTTGGGGTGAATGTTGCGGAAACCATTCGCGGTTCTTTCTACG 4921  
Db 3333 GTTTAAATATATTAATGACAGCACTACTACAGGTGATGTTAACTACTGTTAAAGATCAG 3274  
Qy 4922 CGGTAGAGCTGCTTTTTCGAAATTAATCTTCTTCAATTAATAAACTCTGCTTTGATTACG 4981  
Db 3273 CAGTACAAAAGTTCAACAACCTTCATGCAAAATCTGTTAAGAAACGACAGGTTAAAAAG 3214  
Qy 4982 GAAACGAAGTAAATCCTTTTATGTGAAAGAAATACAAAGTCAATGTACAACTTTGAATG 5041  
Db 3213 AATTAGATCAAGCTGAGCTGATAGAAAAACAAATTAAGACAAACACCAATATGATCAC 3154  
Qy 5042 ATTTCAATTAACAAGCTTTCTGTGAGAGCGCTGCAAGTATTAAGCAGGCTGGAATG 5101  
Db 3153 AACAGAAATTAATGATGCAAAAACAGAAAGTTGATCTGAATTTAAATCAAGGAAAAACA 3094  
Qy 5102 GAGGAATGGTATCTGTCAATCGTGTCTGTATGAAACGGAAGCTTTAGTTAGATTCTG 5161

Db 3093 ATGTCATCATCATCAACAAATGAATATGTTGATTAATGACAGTTAAAGAGAAAAAGCTA 3034  
Qy 5162 AGTTGAAGAGATAGTTCTTTCAATGTAGATGCAAAAAGATCAAAAAACAATTAATACAA 5221  
Db 3033 AAATTAATGCAAGTTAAAAATTTAGTGTAGTACAAAAAGATGCTTTAGCTTAATTAAG 2974  
Qy 5222 TTGCGGAATGCAATGTGAGAAAAAGCGCTGAGATTGAGCAACAGTTGCTCATACAA 5281  
Db 2973 ATGCATTAATGCTTAAGATTAACGAAAGCGGATTAATCTTAACGCACTGCACTTCAAGTAAA 2914  
Qy 5282 ATATTGAAAAACATCACTTATAGCTATTGTAAAAAACAGTAAAAATTCAACGCGGAATG 5341  
Db 2913 TTGCTGAAGCGAAACAAAACCTGCTGATTAATTAACAACTCGCGATCAAAATGTTAATC 2854  
Qy 5342 ATCAAGATAGAAAAAATATCATGTGACTGCAAAAAGATTATACTAGCAATACTATAG 5401  
Db 2853 AAGCTACTTCTTAAGATACATTTGAAGTTCAAAATTCATTAATGACTTAATATTAACG 2794  
Qy 5402 CAGTCGAGTTGAGAGCAAAAGAGCGCTGTGTGCAAGAGCTTGTGCAAGTACTACT 5461  
Db 2793 ATTAACAAATTCACACAGGTAAAAAGAAATCAGCTACACAGATTTATATGCTTATGCA 2734  
Qy 5462 TGAATAGACGTTCTTCTCATGTTGATCAAACTGATATTTGACAAAGATTTAGAGAG 5521  
Db 2733 ATCAGAAAGAAAAATATTTTCAAGTGAACATAATGCAACAGATGAAAAAGCAACAA 2674  
Qy 5522 AAAATTAATGGAATTAAGAAAAAGCAAAATGTTAATGTTCTAGCTGAAATACAGATCAAG 5581  
Db 2673 CAATTAAGCAAGTTGACCAAAATG---TTCAACTGCACTTGAAGCACTTAATATGCTG 2617  
Qy 5582 TGTGTCAAATGCGACAGTGTCTTCGAGAGCAAGTGAACAAGCTGCAAGAGCTGAG 5641  
Db 2616 TGATTAATGTGAGCGTTGATGATGATTAACCAAGGTAAGAGCAGCAATGATGCTATTTC 2557  
Qy 5642 TAGCATTAATTAATTAACAAATACTTGTGCAATTAATAAAATAGTACTCAAAATG 5701  
Db 2556 AAGTAGATGCTACTGTTAACTTAAGGAAACCAAGCTATTTGAAGTTAAACAGAAAGTA 2497  
Qy 5702 TACGAATGCTTTGGTAAAG-----CAATCTCATTCATCTA 5740  
Db 2496 CGAAAGATCTATTGATTAAGATGACCAAGTTAATCTGTGAAGAAAACTGAAGACTTAG 2437  
Qy 5741 TTAACCAATTTGAATGTGAGCTGAGATTGAGCTGAGAGCTGAGAGTGAACAGTTCTG 5800  
Db 2436 CAATGATTAACCAATTAACAGATCAAGCTAAACAAAGGTAATTACTGATCAACCAACACTG 2377  
Qy 5801 TAGCATGAATTAAGATTGTAATTAATACGATAGCAAGATTAAATCATGCAAAAAATCACTG 5860  
Db 2376 CTGAAGTTGAAAAAGCGAAAGCTCAAGGACTTGAAGCATTTGATTAACATTCAAATCGACT 2317  
Qy 5861 CGAAGGGAATGTGGAGTTATTAACAGATCGATGCGGTAAATGCTAATTAATGACAGGA 5920  
Db 2316 CAACGAAAAACAAAAAGCTATCGAAGATTAAGAACTGCACTAGACCAAGATTGAAGCAG 2257  
Qy 5921 CAGTGTCTGAGAGTGGCCGCTGACAGCAATAGAGCTCAACCAAGTGTGAATGAATTAACG 5980  
Db 2256 GTGTAATATGTAAGCTGATGCTACAACTGAAGAAAAAGAGCGTTTACGATGCTTAG 2197  
Qy 5981 GATCTACAAAGCATATGTAAGAAATTTACAGATGATTCTTAAGAGAAAAACAGATGATT 6040  
Db 2196 AAGACATTTTAATCAAAAGCAACGAAAGATTTTCTGATCAAACTACAAATACCAAGAAATCG 2137  
Qy 6041 ATATTACTACTCAAGGGCAAGTAGTAAGTGTGATTAAGTATTCAAAAATCTTAATA 6100  
Db 2136 CTACTGTCAAAAAATAGTGGCTGTGAACACTTAAGCAACAGATTAATCTGAAAGTTA 2077  
Qy 6101 TTAACGAAGCTTATCACAAAAAAGAAAAATAGTAATTAATAAAAGATTTGTACCAATA 6160  
Db 2076 AGAAAAATGCTTTGGAACCAATCAGAGAAAGGTTTAACACAGCAATAAAGAAATTAATAA 2017  
Qy 6161 GTTACGCTACTCATTA 6178  
Db 2016 ATGCAGATGACAGATGCAT 1999

RESULT 2  
US-10-485-517-42/c  
; Sequence 42, Application US/10485517  
; Publication No. US2005025629A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Sheffield  
; APPLICANT: Biosynex Incorporated  
; APPLICANT: Foster, Simon  
; APPLICANT: Mond, James  
; TITLE OF INVENTION: Antigenic Polypeptides  
; FILE REFERENCE: P100629W0  
; CURRENT APPLICATION NUMBER: US/10/485,517  
; CURRENT FILING DATE: 2004-02-02  
; PRIOR APPLICATION NUMBER: GB 0118825.9  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: GB 0200349.9  
; PRIOR FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 424  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 42  
; LENGTH: 2407  
; TYPE: DNA  
; ORGANISM: *Staphylococcus aureus*  
US-10-485-517-42

Query Match 0.6%; Score 63.2; DB 1; Length 2407;  
Best Local Similarity 43.3%; Pred. No. 0.0089;  
Matches 296; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 5185 AATGTGATGCAAAAGATCAAAAACATTAATACATTCGCCGAATGCAATGAGAGA 5244  
DB 874 AACGCATATATATACAGTAAACAAAGAAATTCAAATCTTCAACTACAGAGAA 815  
QY 5245 AAAGCGCTGGAGTGGAGCAAGTTGCTCATCAAAATTTGGAACAAATCAGTTATA 5304  
DB 814 AAACAAGCTGCAATACAGATTTAGACTTAAAGACAGACAGACAGACAAATCTTGAT 755  
QY 5305 GCTATTGTAAAAACAGTAAATTTACACCGCAATGATCAAGATTAAGAAAAATATCAAT 5364  
DB 754 GCTGCAAAATCAAAAGTGAATGTAACAAGCTTAAGACAAATAGTATGCTGCAATTAAT 695  
QY 5365 GTGACTGCAAAAGATTAATCTATGACCAATCTATAGCACTGCGAGTTGAGAGACAAA 5424  
DB 694 CAAGTTCAGCTGCCCAACTAAGAAATCGATGCAAAACGGAATTCGCTCAAAAAGCA 635  
QY 5425 GAGGCTCTGTGCAAGAGCTTCTGCAAGTACCTGTAATTAAGACAGTTTCTTCAT 5484  
DB 634 AGTGAACGTAAACAGCAATTTGAAGCAATGATGATTGCACTACTGAAGAACAAAGCA 575  
QY 5485 GTTGATCAAACTGATATTGACAAAGATTTAGAGAGAAATTAAGAAATTAAGAAAG 5544  
DB 574 GCGAAAGACAAAGTGAATCAAGCAGTAGTACTGCAAAAGCTGATTAATTAATGCTGCA 515  
QY 5545 GCAAAATGTTAATGTTCTAGCTGAAATTAACAAGTCAAGTGTCACAAAATGCGACAGTCTT 5604  
DB 514 GCAAAACATGATGATGATTAATGCAAAAATCAAAATGAACTCAATCGACGACATTACA 455  
QY 5605 TCCGAGCAAGTGAACAAGCTGCAAGTGAAGCTGAGATGAGTAAATTAATTAATTAACAA 5664  
DB 454 CTGATGCAAAATGTTAAACAGCAGCAAAACAGCAATTTGCAATTAAGTAAAGTCAAGCTCAA 395  
QY 5665 AATACCTCTGCACATATAAAATTAAGTCAAAATGTAAGAAATGCTTTGTTAAAGC 5724  
DB 394 GAAACAGCAATTAATGAAATTAACGCTCAACAACTAAAGAAAGACAGCTGCTAAACAA 335  
QY 5725 AATTCATTCATCTATTAATAAACAATTGGAATTGAGCTGAGTGGAGCTGAGAGAGCT 5784  
DB 334 CAAGTTCAAACTGAAAAACAAAGCTGATGCGCAATAGATGACAGACATACAAATGCG 275  
QY 5785 GAGTGAAGGTTCTGATGACAGTGAATTAAGATTGTAAATTAATGATAGAGCAATTAAT 5844

RESULT 3  
US-10-793-626-3493  
; Sequence 3493, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3493  
; LENGTH: 3543  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-3493

Query Match 0.5%; Score 51.4; DB 1; Length 3543;  
Best Local Similarity 42.8%; Pred. No. 1.2; Mismatches 431; Indels 9; Gaps 1;  
Matches 329; Conservative 0;

QY 5141 AAGCTTTAGTATGATCTGAGTTTGAAGATTAAGTTCTTCAATGATGACAAAAG 5200  
DB 601 AAGCATCAGTGCAGATGTAACAAGTTAAACCAAGCATTTGAGAAATCAATGACACAA 660  
QY 5201 ATCAAAAAACATTAATTAATTCGCGCAATGCAATGAGAGAAAAGCGCTGAGTTG 5260  
DB 661 CTGAATTCMAAGAGAACCAATTAATGGAATGATGATCGACAGAGAGAGAGAGAGAG 720  
QY 5261 GAGCAACAGTTCTCTACTAACAAATTTGAAAACAATCAGTTTAGCTATGTTAAATAACA 5320  
DB 721 AAGCTTTAAACAAGTGTATGATCAATTAATTAATCAATTCATTTATTAATTAATGTTA 780  
QY 5321 GTAAATTAACAAGCGCAATGATCAAGATGAGAAAAATATCAATGATGACGCAAGATT 5380  
DB 781 ATTAATTAACAAGATTTATGATTAATTAAGACAAAACGATGAGCTATTAATTAATCA 840  
QY 5381 ATACTATGACCAATTAATGACAGTTCGAGTTGAGAGACAAAAGAGCTCTGTGCAAG 5440  
DB 841 AACCAATATCAACTATCAACCAACCAAGCAATTT-----AAATGAAATCACTATTCAAC 891  
QY 5441 GAGCTTCGCAAGTACTACTCTGATTAAGACAGTTTCTTCTCATGTTGTCACAACTGATA 5500  
DB 892 TAGACACTCAAGCTGATTTAATTAAGAAATTAATTAAGAGTCTACAGTTGAGAAAAAGCCT 951  
QY 5501 TTGACAAAGATTTAGAGAAAGAAATTAATGAAATTAAGAAAAAGCAATGTTAATGTTT 5560  
DB 952 CGGCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1011  
QY 5561 TAGCTGAAAAATGCAAGTCAAGTGTCAAAATGCGACAGTCTTTCCGAGCAAGTGAC 5620  
DB 1012 AAAACAATGAAAGATTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1071  
QY 5621 AAGCTCAAGTGAAGCTGAGTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTA 5680  
DB 1072 TACCTGTTATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1131  
QY 5681 TAAAAAATAGTACTCAAAATGTAAGAAATGCTTTGTTAAAGCAATCTCATTCATCTTA 5740  
DB 1132 TTGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1191



Query Match	0.5#:	Score 48.4:	DB 1:	Length 3327:
Best Local Similarity	46.2#:	Pred. No.3.9:		
Matches 160:	Conservative	0:	Mismatches 186:	Indels 0:
			Gaps 0:	
Qy	5799	TGTAGCAGTGAATTAAGATTGTGTAATTAATACGATGACGAATTAATATCATGCAAAAATCAC	5858	
Db	1565	TGCAGAACCGAATTAATAACAAATCCTATGATGAAAGCAGTTCAAAATGCTGAGTCTATCAT	1624	
Qy	5859	TGCCAAGGGAATGTCGAGTTATTACAGAGTCTGATGCCGTAATTGCTAATTATGCAAG	5918	
Db	1625	TGCAGGATTAATTAATCCACTATCAATTAAGGTAAATGTTATCAAGTGCAGTCTCAAGAGT	1684	
Qy	5919	AACAGTCTCGAGATGGCCGCTGCAGCAATGAGAGCCCAACGATGGAATGAATTAAC	5978	
Db	1685	AATATCATCTAAAAAATGATTAAGTGGTGTTGAACGATTAGCTCAAGATGAGCAAACTGC	1744	
Qy	5979	AGGATCTTACAAAAGCATATGTAAAGATTCTACAGTAGTTGCTTAAGAGAGAAACAGATGA	6038	
Db	1745	TGGAATTCCTTAATATCATTTTAGATTCATTTAAACACAGCTCAACACACAAAGCGTAGAAAA	1804	
Qy	6039	TTATATTACTACAGGGCAAGTAGATAAAGTGTAGATTAAGTATTCAAAAATCTTAA	6098	
Db	1805	TCAATTAATATGCAACAACTCGATGATTAAGTGGCTGGAATTCATTCGACAAAGCGCAAGC	1864	
Qy	6099	TATTAACGAAGACTTATCACAAAAAAGAAAATAAGTATTAATAAAAA	6144	
Db	1865	ATTAAATGAACGATGAAGCAATTAAAAAGAAAGTATTAAGGATCA	1910	

```

RESULT 7
US-10-793-626-4149/C
; Sequence 4149, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUI4980US
; CURRENT APPLICATION NUMBER: US/10/793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4149
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-4149

```

	Query Match	Similarity	0.5%	Score	47.8	DB 1	Length	3543
	Best Local	Similarity	44.9%	Pred. No. 5				
	Matches	267	Conservative	0	Mismatches	322	Indels	6
							Gaps	2
Qy	500	AAGTTTAAATGGAATTCATTCAGTAGATGAAAAGTAAATAATTCATTGAATCCGAATG	559					
Db	1176	AAGCTATACATGCAATTAATGAATTAATGCATAATGAACATTAAGACAAAGATGCCATTA	1117					
Qy	560	GAAGCATTAACGCTAAGAAAATCAATGCTGTGAAGGCAATGCGTTATATGCGGGCG	619					
Db	1116	ATATTTGACTATATCTAGCTGAAAGTAAAAATTCAGATATAGACCAATCAAGTGCAA	1055					
Qy	620	ATATTAGATTGAAAGATACCTGCAATCTAAAGACAGAAATTCACAGATTTTAAAAATTAG	679					
Db	1056	CTACTGAAGAGAAAATACGGCAATACATCTATAGATGATACCTTAGACAAAGACGTA	997					
Qy	680	TCAATTTATAGTATGCA--ATAAATTCTGCTGACCGGAGATTTAAAGTACCAGA	736					
Db	996	ACAATATTATATGCTGCAATATACAAATGCGTTAGTGGATAGCAATTTAGAAAGTGGTAAAC	937					

Qy	737	CAAAATCGGAGATTAATTCTTTCAGTCGATAGATTCCTCCAAAGAACTATGGAA	796
Db	936	AAAGTTACACGTAATGTGTTGTCMACTCAACTAAACACCAAGCTAAAGACAGATTG	877
Qy	797	AAATTCACCTGTTGSAAGAAAGAAATAGAGAAATGTGTAAAGAAATACCAAGCAATA	856
Db	876	CTCAGCAATAGCTTCACCAAGGTCGCACATAGACAGAAATCAAAATGCTTCAACAGAG	817
Qy	857	TTGAATCTGATGCTGTATTGGACGATGAGAAATATTAATAATTAAGTCGAAAGCTACA	916
Db	816	AAAAACAAGAACCCCTTGAGAGACTTAATCAAGAAACAATGAGATCAATGATGAATAC	757
Qy	917	ATGGAGATTTATTAAGAAAGAGGGAAAAAGAACTTATTAACCTCTTTAAGTTAT	976
Db	756	AAGCAGCTTTTGCAAAATCAAAATGTTACAGACGAAAAAAATTAATTAATAGAAACAATA	697
Qy	977	CAGATGTGGAAGCTTCGGTAGAGTAATTAAGGAAAAAGTCATAGGAAGATGTGACA	1036
Db	696	GAAATGTTGAACCTATTTGTAATTGTA---AAACCAAGGCTTAATGAAATTAATTAGAAAAA	640
Qy	1037	TTACAGCTGAAGCAAGAATTTCTATGATGCAACTTTAGTTACTAAGCTTGCAA	1091
Db	639	AAAGTCGGAAACAACGACTTTAATTAATCAAAATCAAGATGCGCACACTAGAGA	585

```

RESULT 8
US-10-793-626-3187
/
/ Sequence 3187, Application US/10793626
/ Publication No. US20050255478A1
/
/ GENERAL INFORMATION:
/ APPLICANT: KIMBERLY, WILLIAM JOHN
/ TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
/ FILE REFERENCE: PUJ4980US
/
/ CURRENT APPLICATION NUMBER: US/10/793,626
/ CURRENT FILING DATE: 2004-03-04
/ PRIOR APPLICATION NUMBER: 60/164,258
/ PRIOR FILING DATE: 1999-11-09
/ NUMBER OF SEQ ID NOS: 4472
/
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3187
/
/ LENGTH: 3840
/
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: synthetic
/
/ OTHER INFORMATION: nucleic acid sequence
/
US-10-793-626-3187

```

Query Match	0.5%	Score 47.8	DB 1	Length 3840
Best Local Similarity	44.9%	Pred. No. 5.1		
Matches	267	Conservative	0	Mismatches 322; Indels 6; Gaps 2;
Qy	500	AAGTTTTTAAATGATCATTCACGTAGATGMAAAGTMAAAATTCATTGATCCGAATG	559	
Db	2039	AAAGCTATCACTGCATTTAAATTGATTAATGCAAAATGCAATMAAAGACAAGATGCATTA	2098	
Qy	560	GAAGCATTACCGTAGAAGAAAAATCAAATGCTGTTGAAGCATCGGTTTATATGCGCGG	619	
Db	2099	ATATTTTGACTTAATCTAGCTGAAGTAAAAATCAGATATTAAGCCAAATCAAGATGCAA	2158	
Qy	620	ATATTAGATTGAAGAATCTGCAATCTMAAAGACAGAAATTACAGATTTTMAAAATTGAG	679	
Db	2159	CTACTGAAGAGAAAAATACGGCAATACATCTTAATGATGATACGTTGACACAAGACGTA	2218	
Qy	680	TCATATTATAGATGCGA---ATAAATTTGGTGTGACCGGAGATTTAAAGCTACCAAGA	736	
Db	2219	ACAAATTTAAATGCTCCAAATACAAATGCGTTAGTGTGATGAGAAATTTAGAAATGCTAAC	2278	
Qy	737	CAAAATCTGAGATATTAATCTTTCAAGTCCACATAGATTCTCTCAAAAGCTATGGGAA	796	
Db	2279	AAAATTTACAACGTATTTGTGTTGCTCAACTCAAACTMAAACACAAGCTAAAGACAGACTTG	2338	
Qy	797	AAAATTCACATGTTGGAAAAGAAATAGAAAGATATGTAAAAAGAAATCCAAAGCAATA	856	

Db 2339 CTCACGACATAGGTCAACAAAGTCGACATAGACAGCAATCAAAATGTACACAGAG 2398  
QY 857 TTGAATTCGATCTGTGATTTGAAGCAGATGAAATTAATAATTTGTGGAAGCTACA 916  
Db 2399 AAAACACAGAGCCCTTGAGAGCTTTAACTCAAGAAACAAATGAGTCAATGATAGAAATC 2458  
QY 917 ATGGAGATTTTAAAGAAAGAGGAGAAAGAAACCTTAACACTCCTTTAAGTTAT 976  
Db 2459 AAGCAGCTTTAGCAATCAAAATGTTACAGAGAAATAATATTTAGAAACATTA 2518  
QY 977 CAGATGTGAAGCTTCCTGAAGATTAATTAAGAAAGAAAGTCTAGGAAAGATTTTGACA 1036  
Db 2519 GAAAGTGTGAACCTATTTGTAATTTGTA---AAACCAAGGCTATGAAATTAATAGAAAA 2575  
QY 1037 TTACAGCTGAGCAAGAAATTTCTATGATGCACTTACTAGCTTGCAAA 1091  
Db 2576 AAGCTGCGGACAAACGACTTTAATTAATCAAAATCAAGATGACACTAGAGAA 2630

RESULT 9  
US-11-074-176-131

/ Sequence 131, Application US/11074176  
/ Publication No. US20050250135A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Kienhammer, Todd R.  
/ APPLICANT: Russell, William M.  
/ APPLICANT: Altermann, Eric  
/ APPLICANT: McAniff, Olivia  
/ APPLICANT: Perill, Andrea Azcarate  
/ TITLE OF INVENTION: Nucleic Acid Sequences Encoding  
/ FILE REFERENCE: 5051-694  
/ CURRENT APPLICATION NUMBER: US/11/074,176  
/ PRIOR FILING DATE: 2005-03-07  
/ PRIOR APPLICATION NUMBER: 60/551,161  
/ NUMBER OF SEQ ID NOS: 381  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 131  
/ LENGTH: 1293  
/ TYPE: DNA  
/ ORGANISM: Lactobacillus acidophilus  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: (1)...(1293)  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: (0)...(0)  
/ OTHER INFORMATION: ORF 1295; Ftey - Cell division protein  
US-11-074-176-131

Query Match 0.5%; Score 45.6; DB 7; Length 1293;  
Best Local Similarity 46.0%; Pred. No. 8.9; Indels 13; Gaps 4;  
Matches 308; Conservative 0; Mismatches 349;

QY 5502 TGACAAAGATTGAGAGAGAAATTAATGAAATTAAGAAAGCAATGTTAATGTTCT 5561  
Db 39 TAAAAATGATGATAT 98  
QY 5562 AGCTGAAATACGAGTCAAGTGTGTCACAAATGCGACAGTCTTCCGAGCAAGTGACA 5621  
Db 99 AAGAT 158  
QY 5622 AGCTGAGTAGAGAGCTGAGAGTACAGTAAATTAATTAACAAATTAATTTGACATAT 5681  
Db 159 TAATGATCTCAAGATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 218  
QY 5682 AAAAAATAGTACCAAAATGTACGAAATGCTTTGTAAGAGAGAGAGAGAGAGAGAT 5741  
Db 219 ATCGGTGAT 278  
QY 5742 TAAAAATGAT 5801

Db 279 AGAAACAGTCTCTGAAGAGTTGTGATGAGAGTCAATCTGAGAGAGAGAGAGAGAGAG 338  
QY 5802 AGCAGTGAATAAGATTGAAATTAATAGATAGCAGAAATTAATCAATGCAAAATCAGTC 5861  
Db 339 ATTAGAGCCGAT 395  
QY 5862 GAAAGAT 5921  
Db 396 AAGAGCAATTAAGAT 452  
QY 5922 AGTGTCTGAT 5981  
Db 453 AGTCATGAT 512  
QY 5982 ATCTACAAAGAT 6041  
Db 513 TGAACAGAT 566  
QY 6042 TATTACTCTCAAGAT 6101  
Db 567 TGCTAAGTCAAGAT 626  
QY 6102 TAAGAT 6161  
Db 627 TAAAGGTGAT 685  
QY 6162 TTACAGTACT 6171  
Db 686 TTACTTAT 695

RESULT 10  
US-10-485-517-11

/ Sequence 11, Application US/10485517  
/ Publication No. US20050256299A1  
/ GENERAL INFORMATION:  
/ APPLICANT: University of Sheffield  
/ APPLICANT: Biosynex Incorporated  
/ APPLICANT: Foerster, Simon  
/ APPLICANT: Mond, James  
/ TITLE OF INVENTION: Antigenic Polypeptides  
/ FILE REFERENCE: P100629W0  
/ CURRENT APPLICATION NUMBER: US/10/485,517  
/ CURRENT FILING DATE: 2004-02-02  
/ PRIOR APPLICATION NUMBER: GB 0118825.9  
/ PRIOR FILING DATE: 2001-08-02  
/ PRIOR APPLICATION NUMBER: GB 0200349.9  
/ PRIOR FILING DATE: 2002-01-09  
/ NUMBER OF SEQ ID NOS: 424  
/ SOFTWARE: PatentIn version 3.1  
/ SEQ ID NO 11  
/ LENGTH: 2275  
/ TYPE: DNA  
/ ORGANISM: Staphylococcus aureus  
US-10-485-517-11

Query Match 0.5%; Score 45.6; DB 1; Length 2275;  
Best Local Similarity 43.4%; Pred. No. 11;  
Matches 210; Conservative 0; Mismatches 274; Indels 0; Gaps 0;

QY 7524 ATCAAGAT 7583  
Db 786 AGCAGTAAATGAT 845  
QY 7584 TTTGTGAT 7643  
Db 846 AGCAGTAAATGAT 905  
QY 7644 AAGAGATATCAAGTCTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 7703  
Db 906 GAAATTAATGAT 965



US-10-793-626-3712

Query Match 0.5%; Score 45.4; DB 1; Length 3180;  
Best Local Similarity 49.0%; Pred. No. 13;  
Matches 121; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 7119 AAATGTAACGGCAAAAGAGAAAAAATATTAAGCAAAACAGCAGCAAAATGCAAGATCGG 7178  
DB 30 AGATCAAAAAGTAGAGCTTGTAAATTAATGTTAAGAAAAATTAGTGAAGCTACTACACAAG 89  
QY 7179 AGGACCAAGTAATCGGACCAATGCTTGTGTAATTAATTTTGAACAAGCTGTGAAGATAG 7238  
DB 90 TGAATTTCAATAGCTGAGCTGATATGCTTATGCAATGATTGGACCGGAAAGA 149  
QY 7239 AAAAAATTCGTAAGAAAAAGAAAGAAAGTTTAACTTGAAGCAAGTTAACAAAGA 7298  
DB 150 TAAAGATTGGAAGAAATTTATGAAAGCGCTTTTGAATATTCACGATATAGCTATTAAAGA 209  
QY 7299 ACAAGATATAAAAGTAATGATGCTACGAAAAAAATCTTACAATCAGCAGGTAATTTCTAC 7358  
DB 210 AGGAGAAACAACTTTTCTGTCATCAATAGCAGAAAGATTAAATTAAGCAGTTGCACTTGGC 269  
QY 7359 AGAAGAT 7365  
DB 270 CAATCAT 276

RESULT 14

US-10-793-626-3756/c  
; Sequence 3756; Application US/10793626  
; Publication No. US2005025478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3756  
; LENGTH: 3650  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3756

Query Match 0.5%; Score 45.4; DB 1; Length 3650;  
Best Local Similarity 49.0%; Pred. No. 13;  
Matches 121; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 7119 AAATGTAACGGCAAAAGAGAAAAAATATTAAGCAAAACAGCAGCAAAATGCAAGATCGG 7178  
DB 1975 AGATCAAAAAGTAGAGCTTGTAAATTAATGTTAAGAAAAATTAGTGAAGCTACTACACAAG 1916  
QY 7179 AGGACCAAGTAATCGGACCAATGCTTGTGTAATTAATTTTGAACAAGCTGTGAAGATAG 7238  
DB 1915 TGAATTTCAATAGCTGAGCTGATATGCTTATGCAATGATTGGACCGGAAAGA 1856  
QY 7239 AAAAAATTCGTAAGAAAAAGAAAGAAAGTTTAACTTGAAGCAAGTTAACAAAGA 7298  
DB 1855 TAAAGATTGGAAGAAATTTATGAAAGCGCTTTTGAATATTCACGATATAGCTATTAAAGA 1796  
QY 7299 ACAAGATATAAAAGTAATGATGCTACGAAAAAAATCTTACAATCAGCAGGTAATTTCTAC 7358  
DB 1795 AGGAGAAACAACTTTTCTGTCATCAATAGCAGAAAGATTAAATTAAGCAGTTGCACTTGGC 1736  
QY 7359 AGAAGAT 7365  
DB 1735 CAATCAT 1729

RESULT 15

US-10-793-626-2229  
; Sequence 2229; Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2229  
; LENGTH: 570  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-2229

Query Match 0.5%; Score 45.2; DB 1; Length 570;  
Best Local Similarity 47.3%; Pred. No. 8;  
Matches 174; Conservative 0; Mismatches 188; Indels 6; Gaps 1;

QY 5862 GAAGGAAATGTCGAGTTATTAACAGACTGATGCGGTAATGCTAATTAAGCAAGAAC 5921  
DB 72 GAAAGACCGATATTAATGATGTTATTAAGCTCTGCGGTAAATAATGATGTTAATTAAGTAAA 131  
QY 5922 AGTGTCTGAGTGGCCCTGAGCAATAGAGCCCTCAACAGTGGATGAATGAATTAACAG 5981  
DB 132 AGACATGCTGTGTAAGCAAGACATAGGTGTACTAGCTTGTGCAAGCAAGAAAGACAT 191  
QY 5982 ATCTCAAAAGCATATGTAAGAAATTTCAAGTG-----ATTGCTAAAGAAAGAAACAGA 6035  
DB 192 AGCTCAAAAGCACTGTATTAAGCTAAGAGATGTGTCTAGCACTGTATTACAAGGATTTGG 251  
QY 6036 TGAATTAATTACTACTCAAGGCAAGTAGATTAAGTGTGTAGATTAAGTATTCAAAATCT 6095  
DB 252 AGATGTTTTGATATGATGAGGTCACTCATGAAAATGTTAATAAAGCTTTGAGAAAGT 311  
QY 6096 TAATTTTAAGCAAGCTTATCACAAAAAGAAAAATAGTAATTAAGAAAGATTGTTAC 6155  
DB 312 TGGTTTAACTTAAGCTTATGAAAAATGACCAATTCATTTGATTAATGACAGCTAT 371  
QY 6156 CAATGATTCAGCTCTACTACTTAAATCTTTAATGCAATGCGCGCTGTTCAAGACA 6215  
DB 372 GATTAAAGAACTTAAAAATGTTAATTAAGACTTTCTTAATGAAGTTTAAGACTGTGACAG 431  
QY 6216 AGCGGAG 6223  
DB 432 CGGTGAG 439

RESULT 16

US-10-793-626-3777/c  
; Sequence 3777; Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3777

LENGTH: 2858  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-3777

Query Match 0.5%; Score 45.2; DB 1; Length 2858;  
Best Local Similarity 47.3%; Pred. No. 13;  
Matches 174; Conservative 0; Mismatches 188; Indels 6; Gaps 1;

QY 5862 GAAGGAAATGCGAGTTATTACAGAGCTGATCGCGTAATGCTAATTAGCAGGAC 5921  
DB 2138 GAAAGACGGTATATTAAGATGTTATAGCTCTGCTGTAATAAATGATGTTAATAAGTAA 2079  
QY 5922 AGTGTGAGAGTGGCCCGGCGAGCAATAGAGCCTCAACCACTGTAAGTAATTAACAG 5981  
DB 2078 AGACATTGGTGTAAAGCAAGACATAGGTGTACTAGCTTTGACAAAGCAAAAGACAT 2019  
QY 5982 ATCTACAAAGCATATGTAAAGATTTCTACAGTG-----ATTGCTAAAGAAAGAAACAG 6035  
DB 2018 AGGTCAAAAGCACTTGAATTAAGCTAAGATGTCTAGCACTGTTATCAAGGTATTGG 1959  
QY 6036 TGATTAATTAATCTACTCAAGGCGAAGTAAAGTGTAGTAAAGTATTCAAAATCT 6095  
DB 1958 AGATGTTTTGTATTTGTAGTCTATCTATGAAATTTGTAATTAAGTCTTTGAGAAAGT 1899  
QY 6096 TAAATTAACGAAGCTTATCACAAAAAGAAATTAAGTAATTAATAAGATTTGTTAC 6155  
DB 1898 TGGTTTAACTTACAGCTTTATGAAAATGACCACTTACATTTGATTAATGACACTAT 1839  
QY 6156 CAATAGTTCAGTACTCATCTTAATAATCTTATTTGGCAATGCGCTGTTGAGAGACA 6215  
DB 1838 GATTAAGAACTTAAATAAGTATTAAGACTTCTTAATGAAGTTTAGACTCTGACAG 1779  
QY 6216 AGCCGGAG 6223  
DB 1778 CCGTGGAG 1771

RESULT 17  
US-10-793-626-3968/c  
Sequence 3968, Application US/10793626  
Publication No. US20050255478A1  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P03480US  
CURRENT APPLICATION NUMBER: US/10/793,626  
CURRENT FILING DATE: 2004-03-04  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3968  
LENGTH: 3002  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-3968

Query Match 0.5%; Score 45.2; DB 1; Length 3002;  
Best Local Similarity 48.4%; Pred. No. 14;  
Matches 156; Conservative 0; Mismatches 163; Indels 3; Gaps 1;  
QY 7686 TGCGTATGCAATTAATTTCTAATGAGATCAATATCAGATTAAATAATTTCTAAGCT 7745  
DB 1176 TACAAAGATGATTAAGAAATTTTCAATGATGATTAATTAATTAATTAATTAATTAAT 1117  
QY 7746 ATTAGAAAAAATATTGATTTTGTAAAAAGTAAATCGAATTGAGACGGAAGCAAA 7805

DB 1116 TTCAGAGCAAAATATTATATTGTTGTACTGATTAACGAAAAGAAAAATTCAGAGAGA 1057  
QY 7806 AGATTTAACCGTAGAGCGGTAGCTGCCGAGCAATATCTCA--AAGCAAGAAATGA 7862  
DB 1056 AGGATATGATCCTGTAATGTGTGTAGACCACTCATTTAGCAAAATTCAAAAGCGTTGA 997  
QY 7863 AATGAATTCAGAGTTGAAATTTGAGACAGATGATTTTCAATGAAGAAATTAAGTAATTAG 7922  
DB 996 AGATTAATTTAAGCGAATTTGATTTAGATGAAATTAATTAAGGTAAAGAAATGAACAT 937  
QY 7923 CCTTCTAAGAAATTTGAGAGAAATCAATGTCAAATGCAAAAAGAAAACAGAGTGAC 7982  
DB 936 TGATCATATGTGTAAAGAAATTTAAGTATGATATTTATGAATTAACGCTAATAAAGAAAC 877  
QY 7983 TGCTGAATCTCAGAGAGCTTCT 8004  
DB 876 AACAGATCATATTAATTAAT 855

RESULT 18  
US-10-793-626-3432/c  
Sequence 3432, Application US/10793626  
Publication No. US20050255478A1  
GENERAL INFORMATION:  
APPLICANT: KIMBERLY, WILLIAM JOHN  
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
FILE REFERENCE: P03480US  
CURRENT APPLICATION NUMBER: US/10/793,626  
CURRENT FILING DATE: 2004-03-04  
PRIOR APPLICATION NUMBER: 60/164,258  
PRIOR FILING DATE: 1999-11-09  
NUMBER OF SEQ ID NOS: 4472  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3432  
LENGTH: 3063  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-3432

Query Match 0.5%; Score 45.2; DB 1; Length 3063;  
Best Local Similarity 47.3%; Pred. No. 14;  
Matches 174; Conservative 0; Mismatches 188; Indels 6; Gaps 1;

QY 5862 GAAGGAAATGCGAGTTATTACAGAGCTGATCGCGTAATGCTAATTATGACGAGAC 5921  
DB 1889 GAAAGACGGTATATTAAGATGTTATTAAGCTCTGCTTAATAAATGATGTTAATTAAGTAA 1830  
QY 5922 AGTGTGAGAGTGGCCCGGCGAGCAATAGGAGCTCAACCACTGTAAGTAATTAACAG 5981  
DB 1829 AGACATTGGTGTAAAGCAAGACATAGGTGTACTAGCTTTGACAAAGCAAAAGACAT 1770  
QY 5982 ATCTACAAAGCATATGTAAAGATTTCTACAGTG-----ATTGCTAAAGAAAGAAACAGA 6035  
DB 1769 AGGTACAAAGCACTTGAATTAAGCTAAGATGTCTAGCACCTGTTATCAAGGTAATTGG 1710  
QY 6036 TGATTAATTAATCTACTCAAGGCGAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 6095  
DB 1709 AGATGTTTTGTATTTGTAGTCTATCTATGAAATTTGGTAATTAAGTCTTTGAGAAAGT 1650  
QY 6096 TAAATTAACGAAGCTTATCACAAAAAGAAATTAAGTAATTAATAAAGATTTGTTAC 6155  
DB 1649 TGGTTTAACTTACAGCTTTATGAAAATGACCATTAATGATTTAATGACAGCTAT 1590  
QY 6156 CAATAGTTCAGTACTCATCTTAAATCTTATTTGGAATTTGCGGCTGTTGAGAGACA 6215  
DB 1589 GATTAAGAAACTTAAATAAGTATTAAGACTTCTTAATGAAGTTTAGACTCTGACAG 1530  
QY 6216 AGCCGGAG 6223

Db 1529 CGGTGAG 1522

## RESULT 19

US-10-793-626-3676/C

Sequence 3676; Application US/10793626  
Publication No. US2005025478A1

## GENERAL INFORMATION:

APPLICANT: KIMBERLY, WILLIAM JOHN

TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

FILE REFERENCE: PUI480US

CURRENT FILING DATE: 2004-03-04

PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 4472

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3676

LENGTH: 3295

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-3676

Query Match 0.5%; Score 45.2; DB 1; Length 3295;

Best Local Similarity 48.4%; Pred. No. 14;  
Matches 156; Conservative 0; Mismatches 163; Indels 3; Gaps 1;

QY 7666 TGCCTATGAGAAATTAATTTCTAATGGAAGATCAATATCAGTATTAAATTTCTAAGCT 7745

DB 1325 TACAAAGATGATTAATAAGAAATGTTACATGATGTAATTAACCTACTACCGCTCT 1266

QY 7746 ATTAGAAAAATATTGATGTTATGTAATAAGATTAATCGGAATTGAGCGGAAGCAA 7805

DB 1265 TTCAGAGCAAAATTAATTAATTTGTTACTGATTAAGGAAAGAAAAATTTGCAAGAA 1206

QY 7806 AGGATTTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCA--AAGCAAGAATGA 7862

DB 1205 AGGATATGATCCTGAAATATGTGTGCTAGACCACTCATTAAGAGCAATTCAAAAACGTTGA 1146

QY 7863 AATGAATTCAGAGGTTGAAATTTGAGAAGATTTTTCATGGAAGAAATAGAGTAAGT 7922

DB 1145 AGATATATTAAACGCAATGATTTTATGATGGAATTAATTAAGGTAAGAAAGTAACAT 1086

QY 7923 CCCTCTTAAGGAATTTGGAAGGAATCAATGCTCAAGTGGAAGAAAAAGAAAGAGTAC 7982

DB 1085 TGATCATGATGTGTAAGAAATTTAAGTATGATTTATGAAATTTACAGCTAAAAAGAAAC 1026

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

DB 1025 AACGATCATATAATTAATTTAT 1004

QY 7983 TGCTGAATCTCAAGAGCTTCT 8004

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic  
OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-4138

Query Match 0.5%; Score 45.2; DB 1; Length 3379;

Best Local Similarity 47.3%; Pred. No. 14;  
Matches 174; Conservative 0; Mismatches 189; Indels 6; Gaps 1;

QY 5862 GAAGGAAATGCGAGTATTATACAGAGCTGATGCGGTAATTTGCTAATTTAGCAGAAC 5921

DB 1261 GAAAGACGATATATTAGATGTTATTTAGCTCTGCTGTAATAAATGATTTAAAGTAA 1320

QY 5922 AGTGTCTGAGTGGCCCGTGCAGCAATGAGACCTCAACAGTGAATGAATTTACAG 5981

DB 1321 AGACATTGTGTGTAAGCAAGACATGAGTGTACTGCTTGTGCAAGCAAGAAAGCAT 1380

QY 5982 ATCTCAAAAGCATATGTAAGATTTACAGTG----ATTGCTAAAGAAAGAAACGA 6035

DB 1381 AGTACAAAGACACTGTGTAAGCTAAAGATGTGCTAGCACTGTTATCAAGGATTTGG 1440

QY 6036 TGATTATATTACTACTCAAGGCAAGTATTAAGTGTAGATAAGTATTCAAATCT 6095

DB 1441 AGATGTTTTTATATATGATGATCCTATGATAATTTGTAATAAGCTTTGAGAAAGT 1500

QY 6096 TAAATTTACGAGACTTATCACAAAAAAGAAATAAGTATTAATAAGATTTGTTAC 6155

DB 1501 TGGTTTAACTTAACTTATGAAAAAATGACACATTTACATTTGATTAAGACAGTAT 1560

QY 6156 CAATGATTCAGCTACTACTTAAATCTTTATTTGGCAATGCCGCTGTTACAGACA 6215

DB 1561 GATTAAAGAACTTAAATAATGATTAATAAGACTTCTTTAATGAAGTTTGAAGCTGACAG 1620

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628

QY 6216 AGCGGAG 6223

DB 1621 CGGTGAG 1628



US-10-793-626-2905

Query Match 0.5%; Score 45.2; DB 1; Length 4719;  
Best Local Similarity 47.3%; Pred. No. 16;  
Matches 174; Conservative 0; Mismatches 188; Indels 6; Gaps 1;

QY 5862 GAAAGGAAATGTCGAGATTATTCAGAGTCGTGCGGTAATTGCTAATTATGACGGAAC 5921  
DB |||||  
DB 3195 GAAAGCGGTATATATGATGTTATAGCTGTGTGTAATAAATGATGTTAATAAGTAA 3254  
QY 5922 AGGTGTGAGTGGCCCGTGCAGCAATGAGAGCTCAACGAGTGAATGAATAATTACAG 5981  
DB |||||  
DB 3255 AGACATTTGTGTGAAGCAAGACATAGGTGTACTAGCTTGCAGCAAGCAAGACAT 3314  
QY 5982 ATCTCAAAAGCATATGTAAAGATTTCTACAGTG-----ATTGCTAAAGAAAGAAACAG 6035  
DB |||||  
DB 3315 AGGTCAAAAGCACTGTATTAAGCTAAAGATGTGTACACCTGTATCAAGGATTTGG 3374  
QY 6036 TGATTTATTTACTACTCAAGGCGAGTATGAATAAGTGAATAAGTAATTCAAAAATCT 6095  
DB |||||  
DB 3375 AGATGTTTTGATTTATGTAGTCTATCTGAAATGTGAATAAAGTCTTGAGAAAGT 3434  
QY 6096 TAAATTAACGAAGCTTATCAAAAAAGAAAAAATAGTAATAAAAAGATTGTTAC 6155  
DB |||||  
DB 3435 TGGTTTAACTTAGACTTTATGAAAAATGACCACTTACCAATTTGATTTATGACAGCTAT 3494  
QY 6156 CATAGTTCAGCTACTCATCTTAAATCTTTATTTGCAAAATGCCGTGTTCAGAGACA 6215  
DB |||||  
DB 3495 GATTAAGAAACTTAAAAATGTAATTAAGACTTTCTTAATGAAGTTTACACTTGCAGG 3554  
QY 6216 AGCCCGAG 6223  
DB |||||  
DB 3555 CGGTGAG 3562

RESULT 25

US-10-485-517-13  
; Sequence 13, Application US/10485517  
; Publication No. US2005025629A1  
; GENERAL INFORMATION:  
; APPLICANT: University of Sheffield  
; APPLICANT: Bioeremus Incorporated  
; APPLICANT: Foster, Simon  
; APPLICANT: Mond, James  
; TITLE OF INVENTION: Antigenic Polypeptides  
; FILE REFERENCE: P100629MO  
; CURRENT APPLICATION NUMBER: US/10/485,517  
; PRIOR FILING DATE: 2004-02-02  
; PRIOR APPLICATION NUMBER: GB 018825.9  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: GB 0200349.9  
; NUMBER OF SEQ ID NOS: 424  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 2119  
; TYPE: DNA  
; ORGANISM: *Staphylococcus aureus*  
US-10-485-517-13

Query Match 0.5%; Score 45; DB 1; Length 2119;  
Best Local Similarity 44.1%; Pred. No. 13;  
Matches 189; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 648 AAAAGCAGAAATTAAGATTTTAAAAATTTAGCATATTAAGATGCAATTAATTTGG 707  
DB |||||  
DB 1686 AACTTACGCTGTTACTCTTAATAAATATGCAATTAATGTTATGAAAAATTTGACAAATC 1745  
QY 708 TCTGACCGGAGATTAAAGTACCAAGCAAAATCTGAGATATTATTTCTTCAGCTCA 767  
DB |||||  
DB 1746 TAAACAGATGCTGCAAGGCAATTTGATGATTTACACATTTTAATATGCAAAAAAGC 1805  
QY 768 CATGATTTCTCTCAAAAAGCTATGGAAAAAATTCACCTGTTGAAAGAGATGAAGA 827

DB 1806 AGATGTTAAATCTAAAAATTAATGCTGCATCAAAATATGCTGCGGTAATTAAGTAAACA 1865  
QY ATATGTTAAAGAAATTAACCAAGCAAAATTAATTAATGATGCTGATGTAATGGAAGAGATGG 887  
DB |||||  
DB 1866 ACAAGTCAAGATTTAAATATACAGGATGGTAACTTGCAGAGTGCAATCAATGATGAACA 1925  
QY 888 AAATATAAAAATTAAGTGGAAAGCTCAAAATGGAGATTTTAAAGAAAGAAAGGAGAAA 947  
DB |||||  
DB 1926 AACGACGGTTAATAGTCAAAACTATCAAGATGCGACACTAGTAAGAAAAACAGCATACAC 1985  
QY 948 AGAAATTAATACATCTCTTTAAGTTATGAGATGTGAAGTTCCTGTAAGATTAATA 1007  
DB |||||  
DB 1986 AAATCGGTACCAAGCTGGCAAGAAATATTTTAAATTAATCAATGTGCTCAAAATTAACGAA 2045  
QY 1008 AGAAAGTCAATGAAAGAAATGTTGACATTTACAGCTGAGAGCAAGAAATTTCTATGATGC 1067  
DB |||||  
DB 2046 AGATCAAGTTACTGAAGGAGATGAATCAAGTGAATTTCTGTAATAAATTAATTAAGATGATC 2105  
QY 1068 AACTTTAGT 1076  
DB |||||  
DB 2106 GCGTTTATT 2114

RESULT 26

US-10-793-626-4448/c  
; Sequence 4448, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; PRIOR FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4448  
; LENGTH: 2621  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-4448

Query Match 0.5%; Score 45; DB 1; Length 2621;  
Best Local Similarity 46.6%; Pred. No. 14;  
Matches 144; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 4373 ATGTCATTGCCAAGATGTCAAGAGAGTTCTGATCTAGCAAAAGAAATATCAAGCTTTAC 4432  
DB |||||  
DB 1956 ATTACATTGATTAAGATGATCAATGAAGAGCAAGATTTTCTTACGCTGTCCAGGTAAAC 1897  
QY 4433 TAAATGAAAGATTAATAAATATTTAGAAAGATCGGTATTAATATGCACTGGAATGGTT 4492  
DB |||||  
DB 1896 CAGAAGATTAAGATTAAGATATTTTGGCGGTCACATATTTATTTATATGATGAATATC 1837  
QY 4493 ATTATGAAAGAACACTAGAAAAAGCAAAAGAAAAAGAAAGAGCGGTCAATTTGAATG 4552  
DB |||||  
DB 1836 ACATACGATTCACACAAAGAAATTTTCGACCTGTGATGCTGTAATCACTTAAATA 1777  
QY 4553 CTGCTTATCGTGTCTGGAACGATTAATCCGCTGAGAGAGTACTACTATGACATTA 4612  
DB |||||  
DB 1776 ATCTTGACGAAGCTATTAATAATTTGCTAATGATTAATTAATGATGCTTCTGCTATGTTA 1717  
QY 4613 CTGTTAAAAATTAATTTTAAAGCAATGAGATGAGCAATTAAGAAACCGGAGAGATA 4672  
DB |||||  
DB 1716 TTGTTAAAGATTAAGATATACATTGAGACATGTTGCTGTTCAATGCAAGCTGTACAAATTG 1657  
QY 4673 AAATTCATG 4681  
DB |||||



QY 674 ATTGATCAATATAGTATCAATTAATCTGCTGACCGAGATTATAAGCTACCA 733  
DB 1144 AATATTCGCTACAGGTGAAGAAATTAAGCTGGATATTTGATACATAAAGTATGC 1203  
QY 734 AGACAAATCTGAGATATATTTCTTTCAGCTCACAATATCTCTCAAAAAGCTATGG 793  
DB 1204 GTATTACTCAGGTGATAGACAGCTTTGATACATACGTTAATAAGTGCGAGTGG 1263  
QY 794 GAAAAAATCAACTGTTGGA 813  
DB 1264 CTCGAAAACCAACTGTTGCA 1283

## RESULT 30

US-10-793-626-4025/c  
; Sequence 4025, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; PRIOR FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4025  
; LENGTH: 3497  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-4025

Query Match 0.5%; Score 44.8; DB 1; Length 3497;  
Best Local Similarity 46.2%; Pred. No. 17;  
Matches 148; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

QY 494 ATGTGTAATGTTTATGATGATCATCTCCAGTATGAGAAAGTAAATTCATTGAATC 553  
DB 3296 AAGTATATGATATATCGGTGAGGTTCCAGTTTCAGAAAAAACAATTGATTAACCAAG 3237  
QY 554 CGAATGGAAGCATTCGCTGAGAGAAAGAAATCAATCTGTTGAAGCATCGGTTATATG 613  
DB 3236 CAATTGGAATATGACTGCTGCTCAAAATCTCTGAGCTATGCGGTAGTAAGTTTG 3177  
QY 614 CGCGGATATGATGATGAAAGATCTGCAATATCTAAAGACAGAAATTACAGATTTTAAA 673  
DB 3176 AACAAACTATGATCTGAAACAACTTTACAAATTGAAATCTTTAAACATTTAGAAA 3117  
QY 674 ATTTAGTCAATATGATGATGATCAATTAATTTCTGCTGACCGAGATTTTAAAGCTACCA 733  
DB 3116 AATATTCGCTACAGGTGAAGAAATTAAGCTGGATATTTGATACATTAAGGTATGC 3057  
QY 734 AGACAAATCTGAGATATATTTCTTTCAGCTCACAATATCTCTCAAAAAGCTATGG 793  
DB 3056 GTATTACTCAGGTGATAGACAGCTTTGATACATACGTTAATAAGTGCGAGTGG 2997  
QY 794 GAAAAAATCAACTGTTGGA 813  
DB 2996 CTCGAAAACCAACTGTTGCA 2977

## RESULT 31

US-10-793-626-3818  
; Sequence 3818, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US

; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3818  
; LENGTH: 3066  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-3818

Query Match 0.5%; Score 44.6; DB 1; Length 3066;  
Best Local Similarity 51.8%; Pred. No. 17;  
Matches 101; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 7183 GCAGCAATCGAGCCATGCTTGTGTAATATATTTTGAACAGCTGTAGAAAGATGAAAA 7242  
DB 2153 GTAGAAACTGAAAGACATCAGAAATGATCAATATATGAAAACTGATCTTACACAGAT 2212  
QY 7243 AATTCTGAAGGAAAGAACAGAACTTTTAAACTTTAGACGAACTTAAACAAGACAA 7302  
DB 2213 AATAGTGAAGCTTAATGCTAGCGATTCAAGAAATAATTCAGAAAGATATCAAGATGAA 2272  
QY 7303 GATAAAGATTAATGATGCTACGAAAAAAATCTTACATGACGAGTATTTCTACAGA 7362  
DB 2273 GAGTCAGAGCTCAAGATCTAATTAATTAAGAAATTAAGAAAGCTTGCAATGACATGAA 2332  
QY 7363 GATACCTCTGTAAAA 7377  
DB 2333 GAAAAATATTTTAAAG 2347

## RESULT 32

US-10-793-626-3849  
; Sequence 3849, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PUS480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; PRIOR FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3849  
; LENGTH: 3170  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-3849

Query Match 0.5%; Score 44.6; DB 1; Length 3170;  
Best Local Similarity 44.6%; Pred. No. 18;  
Matches 271; Conservative 0; Mismatches 324; Indels 12; Gaps 2;

QY 5983 TCTACAAAGCATATGTAAGATTTCTACAGATGATTCCTAAAGAGAAAGATGATAT 6042  
DB 1185 TTTACACATGCAATGGAAGAAATTTAAAGAAATAGCTACAGTACATGATTTTCA 1244  
QY 6043 ATTTACTCAAGGCAAGTATGATTAAGTGTAGATTAAGTATTCAAAAATCTTATATTT 6102  
DB 1245 ATTTAAGATACCAACATCAAGATGATTTGCTTTAAATATTTGCGACATTTGAT 1304  
QY 6103 AACGAAGCTTATCAAAAAAGAAATTAAGTATTAAGATTTGTTACCAATAGT 6162

```
Db 1305 CAACACGATTTGCGCAAAAGTGAAGTCAATTGAACAAAATACTATTAAGT 1364
Qy 6163 TCAGCTACTCATCTTTAAATCTTTATTTGGCAATGCCGCTGTTCAAGACAGCCGA 6222
Db 1365 GGAA-----ATGTTTAAAGCTTACATTAATGACATTTGAGAAATCAATTTGAT 1418
Qy 6223 GTGGAGAGACCTGTTAATATCAACAAGTTTATGGAAGAAAGAAAGCTTTGAGAAAT 6282
Db 1419 AAGGCAATTAATGTAAGTTT-----GGTTAAAGCACTTTTAAACAGTGGCTTTGAAAT 1472
Qy 6283 TCTATATTAATGCAAAACATTAATCTGTAATAATCAGAGATTACAGAAATCAATCGA 6342
Db 1473 GATTAAGTCTTTTGAACAGATTTCAACAATCAGATGATGACCTTGATCGTTAGAA 1532
Qy 6343 GTAGTAGGTTCTGTTGGTGTGGAATGTAGAGATGAGACTTTCTTGATACCAAT 6402
Db 1533 GCACATATTCACAAGAAAGATGAAACAAAGTCAGAGAAAGCACTGAAATAATGAAAAA 1592
Qy 6403 ATTATTAAGAAATATCCAAAGACAGAGTTGGAATACTACATGCTGATGAAGTTTC 6462
Db 1593 ATGAAAGCAGAAAGCGAAACCAAGATTAATTAAGAAAGTACAGTGAATAATGTCAG 1652
Qy 6463 GGAGAAAGAGCTGAATTAACAGAGATTTCTTAAGCAAGAAATTTCTTTTGGAGTCGA 6522
Db 1653 ATTGGAACCAATTCAGATTTGAAATATTAACCAATTTGATCAATTTGAAAGAA 1712
Qy 6523 GTCCAGACGCCGGGTAGAGCCGAGTGCAGAGAACCGTTCCGTAATCAATTTGCA 6582
Db 1713 TTCAAGTAGCTATTAAGGTGTATATTTGATATTAACCTTAAGAACTTAATAAGTGA 1772
Qy 6583 GGAAAGA 6589
Db 1773 CGTCATA 1779
```

## RESULT 33

```
US-10-793-626-2051
; Sequence 2051, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2051
; LENGTH: 4329
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-2051
```

```
Query Match 0.5%; Score 44.6; DB 1; Length 4329;
Best Local Similarity 44.6%; Pred. No. 19;
Matches 271; Conservative 0; Mismatches 324; Indels 12; Gaps 2;
```

```
Qy 5983 TCTACAAAGCATATGTAAGATTTCTACAGTATGCTTAAGAAAGAAAGATGATAT 6042
Db 196 TTACACATGCAATGTAAGAAAGATTTAAAGAAATAGCTACAGTACAGATTTGATTTCA 255
Qy 6043 ATTACTACTCAAGGCGCAAGTATTAAGTGTAGATTAAGTATTTCAAAATCTTAATAT 6102
Db 256 ATTAAGATACCAATCAATCAAGATGAGTGTCTTTAAATAATTTGCGACATTTGATGAT 315
Qy 6103 AACGAAGCTTATCAAAAAAGAAATAATAGTATTAATAAAAGATTTGTTACCAATAGT 6162
Db 316 CAACACGATTTGCGCAAAAGTGAAGTCAATTTGAACAAAATACTATTAAGT 375
```

```
Qy 6163 TCAGCTACTCATCTTTAAATCTTTATTTGGCAATGCCGCTGTTCAAGACAGCCGA 6222
Db 376 GGAA-----ATGTTTAAAGCTTACATTAATGACATTTGAGAAATCAATTTGAT 429
Qy 6223 GTGGAGAGACCTGTTAATATCAACAAGTTTATGGAAGAAAGAAAGCTTTGAGAAAT 6282
Db 430 AAGGCAATTAATGTAAGTTT-----GGTTAAAGCACTTTTAAACAGTGGCTTTGAAAT 483
Qy 6283 TCTATATTAATGCAAAACATTAATCTGTAATAATCAGAGATTACAGAAATCAATCGA 6342
Db 484 GATTAAGTCTTTTGAACAGATTTCAACAATCAGATGATGACCTTGATCGTTAGAA 543
Qy 6343 GTAGTAGGTTCTGTTGGTGTGGAATGTAGAGATGAGACTTTCTTGATACCAAT 6402
Db 544 GCACATATTCACAAGAAAGATGAAACAAAGTCAGAGAAAGCACTGAAATAATGAAAAA 603
Qy 6403 ATTATTAAGAAATATCCAAAGACAGAGTTGGAATACTACATGCTGATGAAGTTTC 6462
Db 604 ATGAAAGCAGAAAGCGAAACCAAGATTAATTAAGAAAGTACAGTGAATAATGTCAG 663
Qy 6463 GGAGAAAGAGCTGAATTAACAGAGATTTCTTAAGCAAGAAATTTCTTTTGGAGTCGA 6522
Db 664 ATTGGAACCAATTCAGATTTGAAATATTAACCAATTTGATCAATTTGAAAGAA 723
Qy 6523 GTCCAGACGCCGGGTAGAGCCGAGTGCAGAGAACCGTTCCGTAATCAATTTGCA 6582
Db 724 TTCAAGTAGCTATTAAGGTGTATATTTGATATTAACCTTAAGAACTTAATAAGTGA 783
Qy 6583 GGAAAGA 6589
Db 784 CGTCATA 790
```

## RESULT 34

```
US-10-793-626-1557/c
; Sequence 1557, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1557
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-1557
```

```
Query Match 0.5%; Score 44.4; DB 1; Length 1500;
Best Local Similarity 52.7%; Pred. No. 15;
Matches 96; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
```

```
Qy 7134 AAGGAAAAAATATTTAAGCAACACAGCAATTCAGAGAAATGCGAGACCAATTCG 7193
Db 699 AGAGAAATTAAGCTTAAGCTTAACATAGACAGAGATGAGAAATATTCAGCTAATTCGG 640
Qy 7194 AGCCAATGCTTGTAAATATTTTGAACAGCTGTAGAGATAGAAAAATTTCTGAGG 7253
Db 639 ACCATTAAACACGTGAAGTAACTCTTGAATACTGTATAAGTACAGAAACCAAGAAAT 580
Qy 7254 AAAAGAACAGAACTTTTAAAACTTTAGACGAAGTTTAAACAAGAACAGATTAAGAT 7313
Db 579 TGAAGGTAAGTTTCTTTAATCTTTTGAACCATCTACAGATTAAGATTAAGTAAGG 520
```

QY 7314 AA 7315  
Db 519 AA 518

## RESULT 35

US-10-793-626-4198/c  
; Sequence 4198, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4198  
; LENGTH: 3145  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-4198

Query Match 0.5%; Score 44.4; DB 1; Length 3145;

Best Local Similarity 45.4%; Pred. No. 19; Indels 0; Gaps 0;  
Matches 159; Conservative 0; Mismatches 191;

QY 4967 CTGCTTTGATTACAGGAACGAAGTAAATCTTTAGTGAAGAATACAAAGTCAATG 5026  
Db 367 CTAGATTATGAAGAAAGTAACTCGGAGAGATTCACAAATTAATGAAATACCTA 308  
QY 5027 TACAAGCTTGAATGATTTCTCATATTTACAAAGCTTTCTGTCGAGCGCTGCAATGTA 5086  
Db 307 AAGATTAATAAATGAAGATTTCTCTTAATATGCTATGTTTAAACAAAAATTTGCTA 248  
QY 5087 AGCAGGCTGAATCGAGGAATGATATGTCATCGTGTCTGATGATAAGGAAGCTT 5146  
Db 247 AGCGTATCCAGTGAATTTGCGCAAGCTATTTGATTAATGATACAAAGCAAGC 188  
QY 5147 TAGTTAGTATCTGAGTTTGAAGAGTAACTTTCTTCAATGATAGTGAAGAAAGTCAA 5206  
Db 187 AGCTTAACAGGTTGATTTGCTGGGCCAGGTTTCATCACTTTATTCGATATCAAT 128  
QY 5207 AAACATAATATCAATTTGCCGGAATGCAATGAGAGAAAGCGCTGAGTTGAGCA 5266  
Db 127 ATTTAACAAGACATTTATTCGAGAAAGCTATCAGAAAGTATGATTTGATGCAAC 68  
QY 5267 CAGTTCATCAATATTTGAAAAATCAATGATTAAGTCTATTGTTAAAA 5316  
Db 67 AATCTAAGATACATAACATTTATTAAGATATGATACGCTAATCCACA 18

## RESULT 36

US-10-793-626-4195/c  
; Sequence 4195, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4195  
; LENGTH: 3328

; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-4195

Query Match 0.5%; Score 44.4; DB 1; Length 3328;  
Best Local Similarity 52.7%; Pred. No. 19;  
Matches 96; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 7134 AGAGAAAAAATATTAGCAACAGCAAGCAATTCGAGAGCAATGCG 7193  
Db 2236 AGAGAAATAGCTATGCTAACATGAGCAAGTATGAAATGATATCAGCTAATTCGG 2237  
QY 7134 AGCAATGCTTGGTAAATTAATTTTGAACAGCTGTAGAGTAAAGAAAAATTCGAAG 7253  
Db 2236 ACCATTAAACACTGTAATATCTTGTAAACGTAATGACAGAAACAAAGAAAT 2177  
QY 7254 AAAAGAACAGAGTTTAAACCTTTAGAGAGTTAAACAAAGAAATTAAGT 7313  
Db 2176 TGAAGTAAAGTTCTTTAATTTCTTTGAACCATCTACATTAAGATTAAGAAAG 2117  
QY 7314 AA 7315  
Db 2116 AA 2115

## RESULT 37

US-10-793-626-3503/c  
; Sequence 3503, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3503  
; LENGTH: 3335  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3503

Query Match 0.5%; Score 44.4; DB 1; Length 3335;  
Best Local Similarity 49.6%; Pred. No. 19; Indels 116; Gaps 0;  
Matches 114; Conservative 0; Mismatches 116;

QY 5943 AGCAATAGAGCCCTCAACCAAGTGAATGAAATTAACAGATCTACAAAAGATATGTAA 6002  
Db 2327 AGCTAACCAACTGTAGATTTTACATTCACGTACAAAAAATAGATGACACTTAAA 2268  
QY 6003 AGATTCAAGTATGCTTAAAGAAAGAAACAGATATTAATTAATCAAGGCAAGT 6062  
Db 2267 ACAAGCAATGTGATTTGACGCGAGTAGCAACCTGTTAGTTACTTAAAGAAATGT 2208  
QY 6063 AGATTAAGTGTAGTAAAGATTTCAAAAATCTTAATTTAAAGAAAGCTTATCAAAA 6122  
Db 2207 CAAAAAGCGCAGTAATTAATGATGTTGTAATACCTATGAAATGAAAAATTTAA 2148  
QY 6123 AAGAAAAATTAAGTAAATAAAGATTTGTTACCAATAGTTCAGTACTC 6172  
Db 2147 AGGAGTATACATATGATGAAGTAAAGAAATAGCTAGTGCATTAATCTC 2098

## RESULT 38

US-10-793-626-4066/C  
; Sequence 4066, Application US/10793626  
; Publication No. US2005025478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4066  
; LENGTH: 3771  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-4066  
Query Match 0.5%; Score 44.4; DB 1; Length 3771;  
Best Local Similarity 45.4%; Pred. No. 20;  
Matches 159; Conservative 0; Mismatches 191; Indels 0; Gaps 0;  
QY 4967 CTGCTTGTATTCACAGAAAGGAAATCCCTTTAGTGAAGAAATACAAAGTCAATG 5026  
DB 2332 CTAGATATTGAAGAAAGTAACTCGCAGAGATATTCCAGAAATTAATAATGAAATACCTA 2273  
QY 5027 TACAAAGCTTGAATGATTTCTATTTACAAAGCTTCTCTGAGAGCGCTGCAAGATTA 5086  
DB 2272 AAGATACCTAAAGAAAGAGATTTCTTTTAATATGCTATGTTTAAACAAAATTTGCTA 2213  
QY 5087 AGCAGGCTGAATCGAGAGATGATGTCTGTCAATGCTGCTCTGATGAACGAAAGCTT 5146  
DB 2212 ACCGTAATCAGTGAATTTGCGCAAGCTATTGTGATCATTTAGATACAGCAAGCAGC 2153  
QY 5147 TAGTATGATTTCTGATTTGAAGAGTAAAGTTCTTTCATATGATGCAAAAGATCAA 5206  
DB 2152 ACGTTAAACAGGTTGATTTGCTGGCCAGGTTTCATCAACTTTATCTGATTAATCAAT 2093  
QY 5207 AAACATAATTAATTCATGCGCGGAAATGCAATGAGAGAAAGCGCTGAGTTGAGCA 5266  
DB 2092 ATTAAACAGACATTTATTCAGAAAGCTATCAAAAGGTATGATTTGATTAACACAC 2033  
QY 5267 CAGTGGCTCATCAATATTGGAAGAAACATCACTTATAGCTATTGTAATA 5316  
DB 2032 AATCTAAGAAATCTACATTTTATTTAGATATGATATCAGCTATCCACA 1983  
RESULT 39  
US-10-793-626-49  
; Sequence 49, Application US/10793626  
; Publication No. US2005025478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 49  
; LENGTH: 2454  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-49

Query Match 0.5%; Score 44.2; DB 1; Length 2454;  
Best Local Similarity 48.6%; Pred. No. 19;  
Matches 152; Conservative 0; Mismatches 158; Indels 3; Gaps 1;  
QY 7686 TGCCATATCAGAAATTAATTTCTAATGAAAGATCAATATCATATTAATAATTTCAAGT 7745  
DB 2142 TACAAAAGATGATTAATAAGAAATTTGTTACATGATGTTAATTAACCTTACACCGCT 2201  
QY 7746 ATTAGAAAAAATATTGATGTTATTGTAAGAAATTAATGCAATTGACAGCGAAACAA 7805  
DB 2202 TTCAGACAAAATATTATATTTGTTACTGATTAACGAAAGAAAAAATTCAGAAAGA 2261  
QY 7806 AGGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCA--AAGCAAGATGA 7862  
DB 2262 AGGATATGATCTCGATATGCTGACCACTCATTTAGAGCAATTCAAAAACGTTGA 2321  
QY 7863 AATGAATTCAGAGTTGAAATTTGAGAAAGATTTTCAATGAAAGAAATGAGTAAGT 7922  
DB 2322 AGATTAATTTAAGCGAATTTGATTTAGATGAAATTAATGAAAGTTAAGAGTAACAT 2381  
QY 7923 CCTTCTAAGGAATTGGAAGAGAAATCAATGTCAAAGTGCAAAAGAAAGAGAGTAC 7982  
DB 2382 TGATCATGATGTTAAGAAATTTAAGTATGATATTATGAAATTTACAGCTAAAAAGAAC 2441  
QY 7983 TGCTGAATCTCA 7995  
DB 2442 AAGCAATCATTA 2454

RESULT 40  
US-10-793-626-1527  
; Sequence 1527, Application US/10793626  
; Publication No. US2005025478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMBERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1527  
; LENGTH: 2454  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-10-793-626-1527  
Query Match 0.5%; Score 44.2; DB 1; Length 2454;  
Best Local Similarity 48.6%; Pred. No. 19;  
Matches 152; Conservative 0; Mismatches 158; Indels 3; Gaps 1;  
QY 7686 TGCCATATCAGAAATTAATTTCTAATGAAAGATCAATATCATATTAATAATTTCAAGT 7745  
DB 2142 TACAAAAGATGATTAATAAGAAATTTGTTACATGATGTTAATTAACCTTACACCGCT 2201  
QY 7746 ATTAGAAAAAATATTGATGTTATTGTAAGAAATTAATGCAATTGACAGCGAAACAA 7805  
DB 2202 TTCAGACAAAATATTATATTTGTTACTGATTAACGAAAGAAAAAATTCAGAAAGA 2261  
QY 7806 AGGATTAACCGTAGAGCGGTAGCTGCCGAGCCATTATCTCA--AAGCAAGATGA 7862  
DB 2262 AGGATATGATCTCGATATGCTGACCACTCATTTAGAGCAATTCAAAAACGTTGA 2321  
QY 7863 AATGAATTCAGAGTTGAAATTTGAGAAAGATTTTCAATGAAAGAAATGAGTAAGT 7922  
DB 2322 AGTATATTTAAGCGAATTTGATTTAGATGAAATTAATGAAAGTTAAGAGTAACAT 2381

Qy	7923	CCCTTCTAAAGGAATTGGAAGAAATCAATGTCAAAGTGA	7982
Db	2382	TGATCATGATGGTAAAGAAATTTAAGTATGATATTATGA	2441
Qy	7983	TGCTGAATCTCAA	7995
Db	2442	AACAGATCATTA	2454

Search completed: November 26, 2005, 06:38:49  
Job time : 317 secs